Heriot-Watt University

Bayesian Analysis of Default and Credit Migration: Latent Factor Models for Event Count and Time-to-Event Data

Yongqiang Bu

June 25, 2014

Submitted for the degree of
Doctor of Philosophy in Mathematical Statistics
on completion of research in the
Department of Actuarial Mathematics and Statistics,
School of Mathematical and Computing Sciences.

I hereby declare that the work presented in this thesis was carried out by myself at Heriot-Watt University, Edinburgh, except where due acknowledgement is made, and has not been submitted for any other degree.

(Candidate)

(Supervisor)

(Supervisor)

Date

Abstract

This thesis develops Bayesian models to explain credit default and migration risk. Credit risk models used in practice are based on an assumption of conditionally independent events given a realization of systematic risk factors. The systematic risk can be modelled with both observed and unobserved factors.

On the one hand we consider generalised linear mixed models (GLMMs) for default count data where random effects account for unobserved factor risk. On the other hand we consider survival models with shared frailties to model unobserved factors in time-to-default and time-to-rating-transition data. The latter models are developed in the Anderson-Gill counting process framework for the Cox proportional hazards model to allow multiple events and time-dependent covariates.

Using Standard and Poor's data on default and rating transitions we control for observed macroeconomic factors in the fixed effect parts of the models. We allow the latent factors to have autoregressive time series structure.

The results from both kinds of model show clear evidence of heterogeneity between industry sectors/countries and time period suggesting that different latent factor effects are present in different sectors. This is an important message that should be accounted for in risk analyses.

We implement Bayesian inference for all our models and use the MCMC approach (Gibbs sampling). We show some tractable model formulations that capture the main sources and implement Bayesian model choice procedures to select the most explanatory models.

There are couple of contributions in this thesis: First, this is an analysis of industry effects on default and migration rates using vector-valued random effects in default count models and vector-valued dynamic frailties in time-to-event/survival models. While this has been done before in models for default counts (McNeil-Wendin) it is quite novel for time-to-event models. Koopman, Lucas and Schwaab (2012) which has some similarities but the estimation is by Monte Carlo maximum likelihood, not by Bayesian methods. Second, estimation of rating transition model with shared dynamic frailties for different industry sectors and macroeconomic covariates using Bayesian techniques (MCMC). This is a new model which is based on a simpler model used in medical statistics (Manda & Mayer(2005)) that has been adapted and extended for the credit risk application. We show how to estimate the new model using a Bayesian approach. Finally, we use the model to compute point-in-time dynamic estimates of rating transition probabilities for different industry sectors and forecast these into the future, while taking into account macroeconomic factors. This can be very useful for risk management applications and economic scenario generation.

Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Alexander McNeil, for introducing me to the world of credit risk modelling, and for all advice and support that I have received during these years. Apart from the always pleasantly informal atmosphere of our research sessions, I especially appreciated his generosity with his time.

I would also like to thank Dr George Streftaris for his comments and advice on Win-Bugs. I am grateful to Prof. Andrew Cairns provide valuable and incomparable comments on my thesis. Thanks also due to all colleagues at Heriot Watt University who made my life at Edinburgh colorful.

Last, but not least, I would like to thank my family for their support and encouragement. Especially to my wife Mingxia and our children.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Credit risk is the risk of the change in value of a portfolio caused by unexpected changes in the credit quality of borrowers, bond issuers or trading partners. Credit risk, together with market risk and operational risk, are the three fundamental risk categories in the banking sector of the financial industry.

The Basel Capital Accord requires banks to hold risk-weighted assets against future potential losses in order to manage the credit risk of business activities. Banks may choose between standardized and advanced approaches for measuring credit risk. In the advanced approach, banks are allowed to build their own internal-ratings-based (IRB) models for measuring certain aspects of credit risk. The IRB approach uses statistical techniques to estimate such quantities as probability of default (PD), loss given default (LGD) and exposure at default (EAD). This has stimulated the development of statistical models in credit risk. Furthermore, the Accord also requires financial institutions to establish rigorous procedures for the validation of statistical models. Thus the measurement and management of credit risk is of interest to both banks and regulators.

Banks need to estimate the credit rating transition probabilities (including default probability) and keep the probability estimation updated for risk management and economic capital purpose. Banks need to compare realized transition probabilities with the estimated probabilities to demonstrate models work reasonably. Those comparisons must use historical data over a long time period. Probabilities of default (PDs) play a prominent role in the areas of pricing credit derivatives, portfolio management and capital allocation. They also determine banks' regulatory capital requirements. For many years, a number of model solutions have been very successful in the financial industry, such as CreditMetrics (CreditManager by RiskMetrics, originally owned by JP Morgan), Portfolio Manager (Moody's KMV), CreditRisk$^+$ (Credit Suisse Financial Products) and CreditPortfolioView (McKinsey). See Crouhy et al. (2000) for a survey. Since the whole industry still has relatively limited historical default data, these credit risk models do not rely on the formal statistical estimation for all parameters from historical default data. Normally we separate the default probability and the parameters which describe the default dependence.

It is well understood that default probabilities partly depend on the macroeconomic situation and rating transitions are also influenced by macroeconomic variables. In building models it is necessary to confirm that this is the case and to find the most explanatory variables. However macroeconomic variables can not fully explain historical patterns of defaults and rating transitions, there is also a need for unobserved (latent) factors in models (McNeil and Wendin (2007)).

All four industrial models above belong to the class of factor models for the approach to dependence (Frey and McNeil (2003)). Gordy (2000) shows these three models share some similarities and can be mapped to each other. While CreditRisk$^+$ belongs to the class of so-called reduced-form models for default modelling, CreditMetrics and Portfolio Manager adopt a structured modelling approach. However these are really issues of presentation and the probabilistic structure of the models is similar. In this thesis, we will describe two different statistical methods for modelling credit risk. These are generalized linear mixed models (GLMMs) which relate quite closely to the industry models described above and survival models with frailty.

Using both factor and reduced-form models, we need to find the most explanatory macroeconomic variables. Several researchers have tried to estimate the rating changes process by finding some relevant explanatory variables. Nickel, Perraudin, and Varotto (2000) estimate a multivariate probit model by working with individual Moodys rating histories. They underline the importance of position in the economic cycle, of individual industry and geographic origin of every underlying firm. Bangia, Diebold, and Shuermann (2002) build quarterly transition matrices with expansions and recessions economic condition. They conclude that the rating transition process can be considered as markovian after conditioning by the state of the economy.

The time-homogeneity property normally assumes that rating migrations are stable over time. However, the time-homogeneity and Markovian behavior assumption have been challenged by many academic studies on the presence of various non-Markovian behaviors such as industry heterogeneity, rating drift and time variations. Several empirical studies have found time-variation in default rates and confirmed that variation could be explained by observed macroeconomic variables; see Nickell, Perraudin, and Varotto (2000), Bangia, Diebold, and Shuermann (2002), and Hu, Kiesel, and Perraudin (2002). The time-varying migration probabilities have become an interesting implications for credit risk modelling when Standard & Poor's rate corporate bonds through the business cycle. Issuer-specific effects in credit risk analysis also become popular in academic research in recent years. McNeil and Wendin (2006) and Koopman, Lucas, and Monteiro (2006) account for issuer-specific effects through latent factors for the business cycle. There is industrial heterogeneity in rating migrations (including default). In this thesis, we will account for both business cycle and industry specific effects by random effects in GLMMs and dynamic frailties in survival models.

There are three contributions in this thesis: First, this is an analysis of industry effects on default and migration rates using vector-valued random effects in default count models and vector-valued dynamic frailties in time-to-event/survival models. While this has been done before in default count models (McNeil and Wendin (2007)), it is quite novel for time-to-event models. Koopman, Lucas and Schwaab (2012) which has some similarities but the estimation is by Monte Carlo maximum likelihood, not by Bayesian methods.

Second, estimation of rating transition model with shared dynamic frailties for different industry sectors and macroeconomic covariates using Bayesian techniques (MCMC). This is a new model which is based on a simpler model used in medical statistics (Manda & Mayer(2005)) that has been adapted and extended for the credit risk application. We show how to estimate the new model using a Bayesian approach. It is very difficult to calibrate the time-to-event model because of the sparsity of data which often leads to unrealistic transition probabilities. Therefore we calibrate the model using Bayesian methods based on Markov Chain Monte Carlo (MCMC) techniques. Bayesian methods improve estimation accuracy especially for low frequency events. Stefanescu, Tunaru, and Turnbull (2009), who also advocate Bayesian methodology for calibrating models for rating transition probabilities using historical data, assert "Model calibration for this type of application is difficult in a classical frequentist estimation framework, because the sparsity of data often leads to unrealistic transition probabilities". Kadam and Lenk (2008) adopt Bayesian estimation techniques for Moody's corporate bond default database and shows strong country and industry effects on the determination of rating migration behavior. Bayesian estimation also allows expert opinion to be taken into account through the use of subjective prior distributions for model parameters. Credit rating process involves a large amount of non-quantifiable subjective information which experienced credit risk practitioners often help to express their opinions. In the case of low default portfolios, expert even gain more weight, especially in industry. Bayesian inference makes it straightforward to compute derived quantities, for example default correlations, the default and transition probabilities, therefore it is increasingly popular; see for instance Nickell, Perraudin, and Varotto (2000), Bangia, Diebold, and Shuermann (2002), Kadam and Lenk (2008) and Stefanescu, Tunaru, and Turnbull (2009). Our model has some attractive features and can be estimated using a standard software package (BUGS), albeit quite slowly. BUGS offers the Deviance Information Criterion (DIC) which developed by Spiegelhalter et al. (2003) to examine the predictive ability of a model.

It is based on a Bayesian measure of predictive power. The model with the smallest DIC value is estimated to give the best predictions for a data set of the same structure as the data actually observed. The DIC measure has the advantage that it does not require the models to be nested for the purpose of comparison.

Finally, we use the model to compute point-in-time dynamic estimates of rating transition probabilities for different industry sectors and forecast these into the future, while taking into account macroeconomic factors. This can be very useful for risk management applications and economic scenario generation.

1.1 Credit risk modelling using GLMMs

Altman (1968) produces an analysis of bankruptcy prediction using Z-scores (*a credit scoring technique*) in his seminal work. This multiple discriminant analysis performs logistic regression using many different accounting ratios. It has become popular among practitioners and provides the basic idea for further regression analyses of default. Merton (1974) assesses the credit risk of a company by characterizing the company's equity as a call option on its assets. He then uses put-call parity to price the value of a put and this is treated as an analogous representation of the firm's credit risk. This model has the limitation that it can only be used for companies with publicly traded equity. For non-listed companies, it is difficult to get asset and liability information.

Credit scoring technique is widely used in banks for building internal rating system in recent years. Banks use probability of default to arrange rating categories. Each rating can be mapped to a probability of default and obligors can be arranged in rating categories using probability of default. The credit rating gives a brief summary of obligors' financial situation, so many banks adopt an internal rating for their own lending management. Merton's model assesses credit risk from asset-liability to model distance to default (DD). Both credit scoring technique and Merton's model allow practitioners to model credit risk. However, not every bank needs to or has the ability to adopt an internal rating system. Therefore third-party rating agencies, such as Moody's, Standard & Poor's and Fitch rating are needed in financial markets. These rating agencies assess the creditworthiness of major publicly listed companies. Servigny and Renault (2004) give a comprehensive discussion of fundamental issues about credit rating.

When considering a portfolio of loans or bonds, the key issue is to model the joint default probability. We must consider the dependence because defaults do not occur independently; knowing the default probability only is far from enough. The dependence between default events has a crucial impact on the upper tail of a credit loss distribution. Default intensities vary over time but share common risk which we refer to as *systematic risk*. Nickell, Perraudin, and Varotto (2000) and Hu, Kiesel, and Perraudin (2002) find time-variation in default rates and confirmed that this time-variation could be explained by observed macroeconomic variables. However, observed variables as proxies for the systematic risk are challenged for the following reasons. First, it is difficult to find appropriate proxies and they usually do not entirely explain the variability in default rates. Second, there may also be a lag between the cycle of a proxy variable and default activity. The lag may vary stochastically over time. These shortcomings can be remedied by latent factors. Koopman, Lucas and Klaassen (2005) give further details about the advantages of latent risk factors. It allow us to capture time-inhomogeneity in default rates and heterogeneity across individual obligors, industry sectors, or any other desired groupings like country with a well-chosen fixed and random effects.

Analytical maximum likelihood techniques can be used for relatively simple models that do not incorporate serial dependence; see for Gordy and Heitfield (2002), Frey and McNeil (2003), and Rosch (2005). McNeil and Wendin (2007) adopt a computational Bayesian methodology with Gibbs sampling with serially correlated random effects. There are some literatures on fitting such models to default data. Crowder et al. (2005) consider a model for default counts with a two-state latent systematic factor following a Markov chain. Gagliardini and Gourieroux (2005), and Koopman, Lucas and Klaassen (2005) consider models where default risk is driven by continuous latent factors using the ratio of defaulted obligors instead of the actual default counts.

Mixture models are also referred as *factor models* or *conditional independence models*. Depending on a set of common economic factors like macroeconomic covariates or latent factors, defaults are assumed to be independent. We will start from the premise of conditionally independent default given a realization of relevant systematic risk factors. The dependence of individual default probabilities comes from the common factors. We will focus on *Bernoulli mixture models* for our analysis; other mixture model like Poisson mixture model can be found in McNeil, Frey, and Embrechts (2005). McNeil and Wendin (2007) highlight the usefulness of *generalized linear mixed models* (GLMMs) in the modelling of portfolio default risk. Both observed and unobserved risk factors can be accommodated by the class of generalized linear mixed models (GLMMs). McNeil and Wendin (2007) and Stefanescu, Tunaru and Turnbull (2009) use Chicago Fed National Activities Index (CFANI) as an important explanatory variable for US firms. There may be a lag between the cycle of macroeconomic variables and that of default activity, we include three months moving average of Chicago Fed National Activities Index (CFANIMA3) in our analysis. And we will compare different macroeconomic covariates in our first set of analyses and then use the best ones in subsequent.

Both McNeil and Wendin (2007) and Stefanescu, Tunaru and Turnbull (2009) use a Bayesian approach to fitting the model with Markov chain Monte Carlo (MCMC) techniques. However there are some nice existing standard software packages which can be used for GLMMs instead of using the complicated Bayesian and MCMC techniques. We use the **glme** function in the S-Plus *correlated data library* for the following analysis. The R function **glmmPQL** is an alternative choice. However, the R function **glmmPQL** can be treated as a special case of **glme** with $(RE)PQL$ method. In S-plus **glme** is a more general function; there are four different methods that can be used in fitting models and these are "AGQUAS","LAPLACE", (restricted) penalized quasi-likelihood ((RE)PQL) and (restricted) marginal quasi-likelihood ((RE)MQL) method. However, (RE)PQL and (RE)MQL gives similar results. Methods AGQUAD and LAPLACE are restricted to family binomial("logit") or poisson("log"). The "probit" link function makes it straightforward to calculate default probability for our one factor model in this analysis, we choose to use the default methods which are "(RE)PQL".

1.2 Credit risk modelling using survival analysis

In the first part of this thesis, we analyse event count data using GLMMs. However, we are not only interested in the number of companies that migrate from one rating category to another, but also interested in the time period the company spends in a certain rating category. Hence, we will study the time-to-event analysis which has become popular in credit risk modelling in recent years. Cox's hazard model has been increasingly used to model the hazard of credit risk events these years. In his seminal work, Cox (1972) proposes the proportional hazards model, where it is possible to estimate the relative intensity of a decrement without specifying the baseline intensity. Cox (1975) demonstrates the estimation procedure for the proportional hazards model with partial likelihood estimation. Andersen and Gill (1982) generalize the model to allow time-varying covariates using a counting process formulation, and show that the maximum partial likelihood estimates are asymptotically equivalent to unconditional maximum likelihood estimates.

As in our count data analysis for GLMMs, we need to capture the time and industry sector heterogeneity using unobserved random variables in survival analysis. The random effects in GLMMs are named frailty in survival framework. Frailty in survival models help to capture heterogeneity with unobserved random variables. We introduce a frailty-based survival model for modelling the intensity of credit rating transitions (default). This type of model is an extension of the Cox proportional hazards models where a common random variable is used to account for heterogeneity. "Frailty models" in survival analysis could capture the unexplained part of the traditional Cox proportional model which take the function of random effects in GLMMs. The frailty factor captures default clustering beyond what can be explained by observed macroeconomic variables and firm-specific information. The unobserved component can capture effects with time or industry sector which show default dependence. Default intensities vary over time and shows co-movements in common or correlated risk factors that all firms are exposed to. Contagion is direct business liaisons between obligors a company may itself face increased risk if one of its major customers defaults. Davis and Lo (2001), and Egloff, Leippold, and Vanini (2007) has investigated this phenomenon.

Kavvathas (2001) and Couderc and Renault (2005) use a similar duration approach conditional on observed macro-variables and they show that average time-to-default decreases as economic activity decreases. Shumway (2001) develops a more dynamic bankruptcy prediction model by combining both financial ratios and market-driven measures and argues that discrete-time is necessary to calibrate hazards because of the intermittency of accounts information. Chava and Jarrow (2004) extend Shumway's (2001) analysis to consider industry sector heterogeneity using monthly intervals. Duffie, Saita, and Wang (2007) formulate a doubly stochastic model for firm survival using firm-specific and macroeconomic covariates.

A credit rating summarises the credit worthiness of an individual, corporation, or even a country. It is an evaluation made by credit bureaus of a borrower's overall credit history. A credit rating is also known as an evaluation of a potential borrower's ability to repay debt, prepared by a credit bureau at the request of the lender. Credit ratings are calculated from financial history and current assets and liabilities. Typically, a credit rating tells a lender or investor the probability of the subject being able to pay back a loan and its interest.

We adapt a simpler model used in medical statistics (Manda and Mayer(2005)) and extended for the credit risk application. We estimate rating transition model with shared dynamic frailties for different industry sectors and macroeconomic covariates using Bayesian techniques (MCMC). This is a model that each transition intensity follows a Cox type multiplicative regression model with frailties, two level frailties to account for both time period and industry sector heterogeneity. Delloye, Fermanian and Sbai (2006) define a reduced-form credit portfolio model which treat rating transition as independent competing risks with conditionally independent and proportional hazards assumption. They also allow strong dependence levels by adding heterogeneity. However, Delloye, Fermanian and Sbai (2006) split the Standard & Poors' data into several groups based on the similar rating transition type and analyze these separately. We consider the whole transition data and let all rating transitions share the same macroeconomic covariates and unobservable random process for all companies in monthly interval.

We consider credit survival model for both default and transition risk and there are two kinds of model were implemented for transition risk. The first one is frailty model for credit rating transitions by numbers of levels (notches) which is a simpler case for credit migration data. It is easy to handle but sacrifice the accuracy for rating transition, therefore we finally model the actual rating transitions. We have shown heterogeneity of transition risk over time and industry sector. We can also show heterogeneity for different countries if we extend our database to all the countries in Creditpro database. For estimation of these models, there are several ways to implement Gibbs simulation. WinBUGS is one of the most popular ways to implement Gibbs simulation. We use the model to compute point-in-time dynamic estimates of rating transition probabilities for different industry sectors and forecast these into the future, while taking into account macroeconomic factors. This is very useful for risk management applications and economic scenario generation.

1.3 Data description

The Standard & Poor's database CreditPro 6.6 which consists of 10439 companies from 13 industry sectors over the period January 1981 to December 2003, 6897 of them are US obligors. The rating classes included in the database are

$$\mathcal{K} = \{CCC, B, BB, BBB, A, AA, AAA, D\}$$

where we merge actual rating k^+, k, k^- into k. We also merge CCC, CC, and C into a single rating class CCC. Rating class AAA and AA which rarely default have been excluded from the default study but be reconsidered in our transition analysis. More than 75% of the companies in the database from US and it is difficult to find macroeconomic covariates to explain credit quality changes for all different countries obligors. The study in this thesis has been restricted to US obligors.

In GLMMs, the default count data have been collected for semester-based periods rating from January 1981 to December 2003 giving a total of $T = 46$ periods. The yearly-based period miss many default events because many obligors migrate many times within one year, therefore we will misreport the default rating, while quarterly data improves the accuracy but too few default events in quarterly period. After comparing these three different data structure, we choose semester-based period in this default study with GLMMs.

In the time-to-event analysis, we choose one month as the time unit. Firstly our macroeconomic covariates are recorded in monthly which match the monthly unit. Secondly, monthly data record most of the transition activities, only very few have transition within one month, the daily would be ideal but it cannot be handled by WinBUGS. Although we use monthly time unit for data manipulation, the time shared frailty is yearly. There are overall 19054 effective rating migrations are recorded in the CreditPro database as well as 1386 defaults. Among them, 13526 effective rating migration as well as 1121 defaults are from US obligors. In time-to-event analysis, we assume the rating migrations only depend on macroeconomic covairates, time period and industry. Without any firm-specific covariates, we only interest in rating migration regardless the obligors. 1.1 shows all the possible transitions in our analysis. We studied default case only in chapter 3 and migration case in chapter 4 with two different models.

	AAA	AA	A	BBB	BB	B	CCC	D	Total
AAA	206	148	14	2	2	0	0	0	372
AA	44	481	553	34	5	5	1	0	1123
A	11	285	1246	840	57	27	1	5	2472
BBB	5	27	521	1342	659	72	9	19	2654
BB	3	8	46	483	1147	883	53	64	2687
B	1	7	25	48	520	1382	850	354	3187
CCC	1	0	4	7	18	129	193	679	1031
Total	271	956	2409	2756	2408	2498	1107	1121	13526

Table 1.1: Numbers of transitions for Standard & Poors' CreditPro 6.6 from 31/12/1980 - 31/12/2003

1.4 Outline of the thesis and main contribution

In this thesis, we model credit risk using two different statistical models which are GLMMs and survival models with frailties. The thesis is structured as follows:

In Chapter 2, we model credit risk using GLMMs and consider the default risk only in this chapter. The Standard & Poor's default count data is used for modelling in this chapter. We allow two levels of heterogeneity for both different times and industries which are represented by random effects in GLMMs framework. We also find three month moving average of Chicago Fed National Activities Index (CFNAIMA3) is the best observed macroeconomic variable to describe the credit default among the macroeconomic variables for US market. We find evidence of significant differences between industry sectors. In Chapter 3, we model default risk using survival model with frailties. We extend the Manda and Meyer (2005) model to allow two levels of random effects and apply this to credit risk modelling for the first time. The Standard & Poor's default time-to-event data is used for modelling in this chapter. We allow two levels of heterogeneity for both different times and industries using survival frailties and serially correlated latent factors. Bayesian inference with Gibbs sampler and WinBugs is used in this chapter. In Chapter 4, we extend the model for default risk to allow multiple events which are ratings transitions. In Chapter 5, we use the intensities results to calculate the credit rating transition probability matrices and give an conclusion in Chapter 6.

Chapter 2

Modelling default risk with GLMMs

Most credit risk models used in practice are based on an assumption of conditionally independent defaults given a realization of systematic risk factors, such models are often referred to as mixture models. The systematic risk can be modelled with both observed factors and unobserved factors. Statistical models from the class of generalized linear mixed models (GLMMs) take observable and unobservable factors as *fixed effects* and *random effects* respectively. Generalized linear model (GLM) is a special case of the GLMM which has no random effects. The random effects in GLMMs help to capture patterns of variability in the response that cannot be explained by the observable factors. The random effects may be scalar or vector. For example, in this thesis both time and industry sector are often treated as two different levels of random effect in default analyses.

There is general lack closed forms for latent factors yields joint default distribution in the form of integrals. Relatively simple models that do not incorporate serial dependence can use analytical maximum likelihood techniques examples include Gordy and Heitfield (2002), Frey and McNeil (2003), and Rosch (2005). McNeil and Wendin (2007) test several models with fixed and random effects using latent factor formulation by Bayesian techniques. Stefanescu, Tunaru and Turnbull (2009) develop a credit rating process model to capture rating transition patterns and estimate it with Bayesian as well. Bayesian methods improve the estimation accuracy especially for low frenquency events. Bayesian estimation also allows expert opinion to be taken through the use of subjective prior distributions for model parameters. Credit rating process involves a large amount of non-quantifiable subjective information which experienced credit risk practitioners often help to express their opinions. In the case of low default portfolios, expert even gain more weight. Bayesian inference becomes straightforward to compute the default and transition probabilities, therefore it is increasingly popular. However, these are a number of software packages for GLMMs based on ML-methodology. The advantage of these defined function is that they are simple and very easy to handle. These defined functions use maximum likelihood inference. We choose to use the **glme** function in the S-Plus *correlated data library* for the following analysis.

In this chapter, we will briefly introduce the mixture models, generalized linear mixed models and GLMMs first. Then we will express the mixture models as GLMMs. We will show how standard software can yield point-in-time(PIT) estimates of default probabilities and illustrate the method using Standard & Poor's CreditPro data but we will provide comparative results with credit risk models using survival models which will be discussed in the next chapter. Alternatively, that may be one or more random effects included in this analyses, with the industry sector random effect, we will investigate differences in default probabilities between industry sectors.

2.1 Theory

2.1.1 One-period mixture models

Default risk is assumed to be driven by systematic risk factors, which might be observed macroeconomic covariates but might also be latent factors. Given a realization of these factors, defaults of individual firms are assumed to be independent. There are several types of mixture models, such as Bernoulli mixture models and Poisson mixture models. More details about these mixture models can be found in McNeil, Frey and Embrechts (2005). Here we consider Bernoulli mixture models in this analysis.

Bernoulli distribution

Consider a portfolio of m obligors. Defaults in a fixed period can be modelled by multivariate Bernoulli distribution. Consider a fixed time period [0,T] and let τ_i be the random time-to-default for obligor $i, i \in \{1, \ldots, m\}$. The default indicator Y_i is a Bernoulli random variable defined by $Y_i = I(\tau_i \leq T)$ so that

$$P(Y_i = 1) = 1 - P(Y_i = 0) = P(\tau_i \leq T) =: PD_i$$

where PD_i is the default probability for obligor i.

Bernoulli mixture model

Give some $p < m$ and a p-dimensional random vector $\mathbf{\Psi} = (\Psi_1, \ldots, \Psi_p)'$, the random vector $\mathbf{Y} = (Y_1, \ldots, Y_m)'$ follows a Bernoulli mixture model with factor vector $\mathbf{\Psi}$, if there are functions $p_i : \mathbb{R}^p \to [0, 1], 1 \leq i \leq m$,such that conditional on $\mathbf{\Psi}$ the default indicator $\mathbf{Y}$ is a vector of independent Bernoulli random variables with $P(Y_i = 1 \mid \mathbf{\Psi} = \psi) = p_i(\psi)$.

For $\mathbf{y} = (y_1, \ldots, y_m)'$ in $\{0, 1\}^m$ we have

$$P(\mathbf{Y} = \mathbf{y} \mid \mathbf{\Psi} = \boldsymbol{\psi}) = \prod_{i=1}^{m} p_i(\psi)^{y_i}(1 - p_i(\psi))^{1-y_i}$$

and the unconditional distribution of the default indicator vector $\mathbf{Y}$ is obtained by integrating over the distribution of the factor vector $\mathbf{\Psi}$.

$$P(Y_1 = y_1, \ldots, Y_n = y_n) = \underset{\mathbb{R}^p}{\int \cdots \int} \prod_{i=1}^{m} p_i(\psi)^{y_i}(1 - p_i(\psi))^{1-y_i} dG(\psi)$$

where $y_1, \ldots, y_n \in \{0, 1\}$,and G is a distribution function on $\mathbb{R}^p$.

Exchangeability and mixture models

To simply the analysis we will often assume that default indicator are exchangeable for obligors in a group. The sequence $Y_1, \ldots, Y_m$ of random variables is said to be exchangeable if

$$(Y_1, ..., Y_m) \stackrel{\mathrm{d}}{=} (Y_{\Pi(1)}, \ldots, Y_{\Pi(m)})$$

for any permutation $(\Pi(1), \ldots, \Pi(m))$ of $1, \ldots, m$. For a group of similarly rated company without any other information, the assumption of exchangeability is stronger than merely assuming identical marginal distribution of $Y_1, \ldots, Y_m$ but weaker than assuming $Y_1, \ldots, Y_m$ to be independent and identically distributed (i.i.d). We introduce a simple notation for default probabilities where

$$\pi := P(Y_i = 1), \quad i \in \{1, \ldots, m\}$$

is the default probability of any obligor and

$$\pi_k := P(Y_{i_1} = 1, \ldots, Y_{i_k} = 1), \quad \{i_1, \ldots, i_k\} \subset \{1, \ldots, m\}, \quad 2 \leq k \leq m$$

is the joint default probability for k firms. When default indicators are exchangeable we get

$$E(Y_i) = E(Y_i^2) = P(Y_i = 1) = \pi, \quad \forall i,$$
$$E(Y_i Y_j) = P(Y_i = 1, Y_j = 1) = \pi_2, \quad \forall i \neq j,$$

so that $\mathrm{cov}(Y_i, Y_j) = \pi_2 - \pi^2$; then default correlation is give by

$$\rho_Y := \rho(Y_i, Y_j) = \frac{\pi_2 - \pi^2}{\pi - \pi^2}, \quad \forall i \neq j \tag{2.1}$$

Exchangeable Bernoulli mixture models

Let m denote the number of observed companies and M denote the number that default. Assume that all the p_i are identical function. Bernoulli mixture model is exchangeable since default indicator $\mathbf{Y}$ is exchangeable. Let us introduce $Q := p_1(\Psi)$ and take the distribution function of Q to be G(q) conditional on $Q = q$. The number of defaults M is the sum of m independent Bernoulli variables with parameter q. It is given by a binomial distribution with parameters q and m.

$$P(M = k \mid Q = q) = \binom{m}{k} q^k (1-q)^{m-k}$$

The unconditional distribution of M is obtained by integrating over q

$$P(M = k) = \int_0^1 \binom{m}{k} q^k (1-q)^{m-k} dG(q)$$

The default probability and joint default probabilities for the exchangeable group are given by:

$$\pi = E(Y_1) = E(E(Y_1 \mid Q)) = E(Q)$$

$$\pi_k = P(Y_1, \ldots, Y_k = 1) = E(E(Y_1, \ldots, Y_k \mid Q)) = E(Q^k)$$

For $i \neq j$

$$\mathrm{cov}(Y_i, Y_j) = \pi_2 - \pi^2 = \mathrm{var}(Q) \geq 0,$$

The default correlation ρ_Y defined in (2.1) for exchangeable Bernoulli mixture model is always non-negative.

Firm-valued models as Bernoulli mixture models

The Merton model is the prototype of all firm-value models. Merton model has been extended over the years but the original remains an influential benchmark and is still popular in credit risk analysis. A firm i whose asset value follows some stochastic process $V_{t,i}$ has one single debt with face value B_i and maturity T. The process $V_{t,i}$ follows a diffusion model under real-world probability measure P

$$dV_{t,i} = \mu V_{t,i} dt + \sigma_{V_i} V_{t,i} dW_t$$

which implies that

$$V_{T,i} = V_{0,i} \exp\left((\mu_{V_i} - \frac{1}{2}\sigma_{V_i}^2)T + \sigma_{V_i} W_{T,i} \right)$$

The default probability of firm i is given by

$$P(V_{T,i} \leq B_i) = P(\ln V_{T,i} \leq \ln B_i) = \Phi\left(\frac{\ln \frac{B_i}{V_{0,i}} - (\mu_{V_i} - \frac{1}{2}\sigma_{V_i}^2)T}{\sigma_{V_i}\sqrt{T}}\right)$$

Industry models can be re-written as mixture models. We assume that default occurs for obligor i if a critical value X_i (asset value, $V_{T,i}$ in Merton's model) lies below a critical threshold d_i (liabilities, B_i in Merton's model) at the end of each time period. KMV/CreditMetrics is industry credit risk models using Merton-type structure. They can be expressed by a mixture model. In order to apply these models at portfolio level require a multivariate Merton's model. Assume that we have m companies and that the multivariate asset-value process $(\boldsymbol{V_t})$ with $\boldsymbol{V_t} = (V_{t,i}, \ldots, V_{t,m})'$ follows an m-dimensional geometric Brownian motion with drift vector $\boldsymbol{\mu_V} = (\mu_{V_i}, \ldots, \mu_{V_m})'$, vector of volatilities $\boldsymbol{\sigma_V} = (\sigma_{V_i}, \ldots, \sigma_{V_m})'$ and instantaneous correlation matrix P. This implies that for any firm i default occurs when some critical rv $X_i := X_{T,i}$ lies below some critical deterministic threshold d_i at the end of the time period $[0, T]$. In Merton's model X_i is a log-normally distributed asset value and d_i represents liabilities. Here we typically use multivariate log-normal or normal distribution for the vector $\boldsymbol{X} = (X_i, \ldots, X_m)'$. The dependence among defaults comes from the dependence among the components of the vector $\boldsymbol{X}$.

We consider a portfolio of m obligors and fix a time horizon T. For $1 \leq i \leq m$, we let rv S_i be a state indicator for obligor i at time T and assume that $S_i := S_{T,i}$ takes integer values in the set $0, 1, \ldots, n$ representing rating class. We interpret 0 as default state. Here we will concentrate on the binary outcomes of default and non-default and ignore the finer categorization of non-defaulted companies. We write $Y_i := Y_{T,i}$ for the default indicator variable so that $Y_i = 1 \Leftrightarrow S_i > 0$. Random vector $\boldsymbol{Y} = (Y_i, \ldots, Y_m)'$ is a vector of default indicators for the portfolio and $p(\boldsymbol{y}) = P(Y_1 = y_1, \ldots Y_m = y_m), \boldsymbol{y} \in (0,1)^m$ is its joint probability function; the marginal default probabilities are denoted by $\bar{p_i} = P(Y_i = 1), i = 1, \ldots, m$. The default correlations are defined to be the correlation of the default indicators. $var(Y_i) = E(Y_i^2) - \bar{p_i}^2 = E(Y_i) - \bar{p_i}^2 = \bar{p_i} - \bar{p_i}^2$, then we obtain for firms i and j, with $i \neq j$,

$$\rho(Y_i, Y_j) = \frac{E(Y_i Y_j) - \bar{p_i}\bar{p_j}}{\sqrt{(\bar{p_i} - \bar{p_i}^2)(\bar{p_j} - \bar{p_j}^2)}} \tag{2.2}$$

It is important to distinguish the default correlation $\rho(Y_i, Y_i)$ of two firms $i \neq j$ from the asset correlation which is the correlation of the critical variables X_i and X_j.

The vector of critical variables $\boldsymbol{X}$ is assumed to have a multivariate normal distribution and X_i can be interpreted as a change in asset value for firm i over the time horizon of interest; d_{i1} is chosen so that the probability that $X_i \leq d_{i1}$ matches the given default probability $\bar{p_i}$ for firm i.

The covariance of X is calibrated using a factor model. We assume that X can be written as

$$\boldsymbol{X} = \boldsymbol{B}\boldsymbol{F} + \boldsymbol{\varepsilon} \tag{2.3}$$

for a p-dimensional random vector of common factors $\boldsymbol{F} \sim N_p(\mathbf{0}, \Omega)$ with $p < m$, a loading matrix $\boldsymbol{B} \in \mathbb{R}^{m \times p}$, and an m-dimensional vector of independent univariate normally distributed errors $\boldsymbol{\varepsilon}$, which are also independent of $\boldsymbol{F}$. The factor structure (2.3) implies that the covariance matrix P of $\boldsymbol{X}$ (which will be a correlation matrix due to our assumptions on the marginal distribution of $\boldsymbol{X}$) is of the for $P = B\Omega B' + \Upsilon$, where Υ is the diagonal covariance matrix of $\boldsymbol{\varepsilon}$.

The conditional independence of defaults given $\boldsymbol{\Psi}$ follows from the independence of the idiosyncratic terms $\varepsilon_1, \ldots, \varepsilon_m$. Take $\boldsymbol{b_i} = (b_{i1}, \ldots, b_{ip})'$ for the ith row of B, the ith critical variable has the following structure:

$$X_i = \mathbf{b}_i'\mathbf{F} + \varepsilon_i \tag{2.4}$$

where $\varepsilon_i \sim N(0, 1 - \beta_i)$ with $\beta_i = \mathrm{var}(\mathbf{b}_i'\mathbf{F}) = \mathbf{b}_i'\Omega\mathbf{b}_i$, independent of $\mathbf{F}$ and of ε_j for $j \neq i$. X_i can represent $\ln V_T$ in Merton's model.

Asset correlation

The three most important drivers in determining portfolio credit risk are default probability(PD), loss given default (LGD) and default correlation. The most common approach to modelling default correlation (defined in equation 2.2)is to combine default probability with asset correlation. Therefore asset correlation is a critical driver in modelling portfolio credit risk. Two obligors will default in the same time if both of their asset value is smaller than their obligations. Asset correlation helps define the joint behavior of the asset value of two obligors. The asset correlations show how the asset value of one obligor depends on another obligor's asset value. The correlation can also be described as the dependence of the asset value of an obligor on the general state of the economy and all obligors are depend on the state of economy which can be described by macroeconomic covariates.

Basel III IRB framework gives a formula for credit risk charge for an exposure using asset correlation ρ which is the average asset correlation of all pair-wise asset correlation in the portfolio. Here we consider the asset correlation using (2.4). The asset correlation between companies i and j is given by

$$\rho(X_i, X_j) = \mathrm{cov}(X_i, X_j) = E(X_i X_j) = \mathbf{b}_i' \Omega \mathbf{b}_j$$

The asset correlation for practical models will be described in the models we use.

2.1.2 Generalized linear mixed models and its estimation

Clayton (1996) gives a very good overview of generalized linear mixed models(GLMMs). The special case of GLMMs without random effect all known as generalized linear models (GLMs). The random effects not only determine the correlation structure between observations on the same group, they also take account of heterogeneity among groups that is attributed to unobserved factors. We will start with generalized linear models and then extend GLMs to GLMMs by adding random effects. Different inference methods for GLMMs will be introduced.

Generalized linear mixed models

Generalized linear models (GLMs) extend the linear model to allow distributions from the exponential family. The outcome of the response variables, $\boldsymbol{Y} = (y_1, \ldots, y_n)'$, is assumed to be generated from a particular distribution function in the exponential family. The mean, $\boldsymbol{\mu} \equiv (\mu_1, \ldots, \mu_n)'$, of the distribution depends on the explanatory variables, $\boldsymbol{X} \in \mathcal{R}^{n \times p}$, through:

$$E(\mathbf{Y}) = \boldsymbol{\mu} = g^{-1}(\boldsymbol{X\beta})$$

where $E(\boldsymbol{Y})$ is the expected value of $\boldsymbol{Y}$, $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_p)$; $\boldsymbol{X\beta}$ is the linear predictor, a linear combination of unknown parameters, $\boldsymbol{\beta}$; g is the link function. The unknown but fixed regression parameters in $\boldsymbol{\beta}$ are estimated by solving the maximum likelihood equations. Generalized linear models (GLMs) consists of three components: A distribution function f from the exponential family; a linear predictor $\boldsymbol{\eta} = \boldsymbol{X\beta}$; and a link function g such that $E(\boldsymbol{Y}) = \boldsymbol{\mu} = g^{-1}(\boldsymbol{\eta})$. The link function provides the relationship between the linear predictor and the mean of the distribution function. There are many commonly used link functions, such as probit and logit link function in the Bernoulli case and log-link function in the Poisson case. Here we give the common response functions (g^{-1}) for Bernoulli-type GLMMs.

- Probit
$$\Phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^{x} \exp(-u^2/2) du$$
- logit
$$e^x/(1+e^x) = 1/(1+e^{-x})$$
- complementary log-log
$$1 - \exp(-e^x)$$

Generalized linear mixed models (GLMMs) are GLMs with one or more random effects. The linear predictor should be rewritten as

$$\boldsymbol{\eta} = \boldsymbol{X\beta} + \boldsymbol{Z\psi}$$

where the fixed effect $\boldsymbol{\beta}$ remains the same as in GLMs and random effects $\boldsymbol{\psi}$ are random variables drawn from a distribution. The random effects generate heterogeneity beyond that which can be captured with fixed effects. We firstly consider a single shared random effect in every time period in our model. Then we consider different random effects for different sectors. This will capture additional variability associated with economic effects in different sectors. $\boldsymbol{X}$ are fixed vector and could be rating and macroeconomic covariates in credit risk models, $\boldsymbol{\beta}$ are the corresponding fixed effects. While $\boldsymbol{Z}$ are random vectors and could be year and industry sectors, $\boldsymbol{\psi}$ are random effects.

Estimation of GLMMs

We can estimate the GLMMs parameters using Markov Chain Monte Carlo(MCMC) methods from a Bayesian point-of-view. Bayesian MCMC approach allows us to handle more complex models than standard packages. Furthermore, the Bayesian approach can work well for sparse default data. For a Bayesian approach to GLMMs of portfolio credit risk, see McNeil and Wendin (2007) for more information. However, some standard package like S-plus and R provide some defined functions for estimation of GLMMs. The advantage of these defined function is that they are simple and very easy to handle. These defined functions use maximum likelihood inference. We will give more details about maximum likelihood inference in the next paragraph.

Maximum likelihood (ML) inference for GLMMs is only really a possible option for the simplest model. The unconditional distribution is obtained by integrating the random effects. Factor models have conditional independence property, if we write $p_{Y_{t,i}|\Psi_t}(y \mid \psi)$ for the conditional probability mass function of $Y_{t,i}$ given Ψ_t,

$$L(\beta, \sigma; data) = \int \cdots \int \left(\prod_{t=1}^{n} \prod_{i=1}^{m_t} p_{Y_{t,i}|\Psi_t}(Y_{t,i} \mid \psi) \right) f(\psi_1, \ldots, \psi_n) d\psi_1 \ldots d\psi_n$$

where f denotes the joint density of the random effects. With the *iid* Gaussian random effects with marginal Gaussian density f_Ψ, we can reduce the $n-$dimensional integral to

$$L(\beta, \sigma; data) = \prod_{t=1}^{n} \left(\int \prod_{i=1}^{m_t} p_{Y_{t,i}|\Psi_t}(Y_{t,i} \mid \psi) f_\Psi(\psi_t) d\psi_t \right)$$

with the product of one-dimensional integrals. This can be easily solved the unknown parameters.

For those complicated models where the exact likelihood function is difficult to compute, approximation becomes unavoidable. There are several methods including penalized quasi-likelihood (PQL), marginal quasi-likelihood(MQL), Laplace and adaptive gaussian quadrature. More details about PQL and MQL can be found in Breslow and Clayton (1993).

We use the **glme** function in the S-Plus *correlated data library* for the following analysis. The R function **glmmPQL** is an alternative choice. However, the R function **glmmPQL** can be treated as a special case of **glme** with $(RE)PQL$ method. In S-plus **glme** is a more general function; there are four different methods that can be used in fitting models and these are "AGQUAS","LAPLACE", (restricted) penalized quasi-likelihood ((RE)PQL) and (restricted) marginal quasi-likelihood ((RE)MQL) method. However, (RE)PQL and (RE)MQL gives similar results. Methods AGQUAD and LAPLACE are restricted to family binomial("logit") or poisson("log"). The "probit" link function makes it straightforward to calculate default probability for our one factor model in this analysis, we choose to use the default methods which are "(RE)PQL".

2.1.3 Multi-period mixture models

Notation

Consider a multi-period model, we can write the general latent variable factor model as

$$X_{it} = \mathbf{b}_{it}'\mathbf{F}_t + \epsilon_{it} \tag{2.5}$$

where $\mathbf{F}_t \sim N_p(\mathbf{0}, \Omega)$ and $\epsilon_{it} \sim N(0, 1 - \beta_{it})$ with $\beta_{it} = \mathrm{var}(\mathbf{b}_{it}'\mathbf{F}_t) = \mathbf{b}_{it}'\Omega\mathbf{b}_{it}$. $F_1, \cdots, F_t$ does not necessary to be *i.i.d.* X_{it} can represent $\ln V_T$ in Merton's model. In the credit risk model default occurs according to an indicator variable $Y_{it} = I_{(X_{it} \leq d_{it})}$

so that the conditional default probability is given by

$$\begin{aligned} P(Y_{it} = 1 \mid \mathbf{F}_t) &= P(\epsilon_{it} \leq d_{it} - \mathbf{b}'_{it}\mathbf{F}_t \mid \mathbf{F}_t) \\ &= \Phi\left(\frac{d_{it}}{\sqrt{1-\beta_{it}}} - \frac{\mathbf{b}'_{it}\mathbf{F}_t}{\sqrt{1-\beta_{it}}}\right) \end{aligned} \tag{2.6}$$

and the unconditional default probability by $p_{it} = \Phi(d_{it})$.

In this simple case of (2.5) where is only one factor we have

$$X_{it} = b_{it}F_t + \sqrt{1 - b_{it}^2}Z_{it}$$

for iid standard normal variables $F_t, Z_{1t}, Z_{2t}, \ldots$ and a common loading $-1 < b_{it} < 1$. The asset correlation between credit risks i and j in time period t is given by $\rho_{ijt} = b_{it}b_{jt}$ and the conditional default probabilities by

$$P(Y_{it} = 1 \mid F_t) = \Phi\left(\frac{d_{it}}{\sqrt{1-b_{it}^2}} - \frac{b_{it}F_t}{\sqrt{1-b_{it}^2}}\right). \tag{2.7}$$

Mixture models as GLMMs

In one-period setting, we assume conditional default probabilities $p_i(\Psi)$ follows

$$p_i(\Psi) = h(\mu + \boldsymbol{\beta}'\boldsymbol{x}_i + \Psi)$$

where h is a link function, such as probit, logit and complementary log-log which are commonly used in GLMMs. The vector $\boldsymbol{x}_i$ contains covariates for company i, such as company specific information or industry, country group $\boldsymbol{\beta}$ and μ are model parameters. The random variable $\Psi \sim N(0, \sigma^2)$ is normally distributed with scale parameter σ.

This model can be extended into multi-period model which is suitable for default counts for different time periods. Consider a series of mixing variables $\Psi_1, \ldots, \Psi_n$ generate default dependence in each time period $t = 1, \ldots, n$. The default indicator $Y_{t,i}$ for company i in time period t is assumed to be Bernoulli with default probability $p_{t,i}(\Psi_t)$ depending on Ψ_t

$$p_{t,i}(\Psi_t) = h(\mu + \mathbf{x}'_{t,i}\beta + \Psi_t)$$

where $\Psi_t \sim N(0, \sigma^2)$ and $\mathbf{x}_{t,i}$ are covariates for company i at time period t.

Dynamic latent effect

Both the one-period and multi-period Bernoulli mixture models belong to GLMMs family. The role of the random effects in GLMMs is to capture patterns of variability in the responses that cannot be explained by the observed covariates only. These non-observed factors could be time-period effect which we treat as the state-of-the-economy in that time period.

The GLMM framework allows more complexity by adding further random effects to obtain multi-factor mixture models. In our analysis, we also consider industry sector effects nested inside the time-period effects which is two-levels random effects. We can capture additional variability in different industry sectors in addition to the global variability given by time-period effect.

2.2 Models used in practice

It is useful to consider a one-factor model in many practical situations. The information may not always be available to calibrate a model with more factors, and one-factor models may be easily fitted statistically to default data. Here we consider the models for default risk with GLMMs.

2.2.1 Model 2.1: One-factor model with Equicorrelation structure

Let us consider the case where the default probability depends only on the credit rating $r(i,t)$ of the ith individual in period t and where loadings b_{it} for all individuals in all time periods are the same so that

$$P(Y_{it} = 1 \mid F_t) = \Phi\left(\frac{\Phi^{-1}(p_{r(i,t)})}{\sqrt{1-b^2}} - \frac{bF_t}{\sqrt{1-b^2}}\right),$$

where b is the common loading and the asset correlation is $\rho = b^2$. This model can be written in simpler GLMM notation as

$$P\left(Y_{it} = 1 \mid \Psi_t\right) = \Phi\left(\gamma_{r(i,t)} + \Psi_t\right)$$

where γ_r is a fixed rating effect and $\Psi_t \sim N(0, \sigma^2)$ is a random effect for period t. The parameters in the credit risk model and the GLMM are related by

$$\begin{aligned} \rho = b^2 &= \frac{\sigma^2}{1+\sigma^2} \\ p_r &= \Phi\left(\frac{\gamma_r}{\sqrt{1+\sigma^2}}\right) \end{aligned}$$

so that both asset correlation and rating class default probabilities may be inferred from the fitted GLMM.

2.2.2 Model 2.2: Model with macroeconomic covariates

The one-factor model with equicorrelation structure is extended by adding observed macroeconomic covariates z_t. We will try to include several different covariates in our analysis. Default probability depend on the credit rating effect $\gamma_{r(i,t)}$ of ith individual in period t and macroeconomic covariates $\boldsymbol{z_t}$ ($\boldsymbol{z_t}$ could be vector or scale). The default probability in the same time period with the same rating category will be the same. This model can be written in simpler GLMM notation as

$$P\left(Y_{it} = 1 \mid \Psi_t\right) = \Phi\left(\gamma_{r(i,t)} + \boldsymbol{\eta} \boldsymbol{z}_t + \Psi_t\right)$$

where $\boldsymbol{\eta}$ is an additional parameter. The implied asset correlation in period t. If we also consider correlation coming from macroeconomic factor we get:

$$\rho = \frac{\sigma^2}{1+\sigma^2}$$

2.2.3 Model 2.3: Model with sector random effects

Now consider the special case of (2.5) where $\mathbf{b}_{it} = b\mathbf{e}_{s(i,t)}$ where $\mathbf{e}_j$ denotes a unit vector and $s(i,t)$ gives sector membership. Assume there are p sectors and that Ω is a symmetric equicorrelation matrix with equicorrelation parameter ρ.

Since $\beta_{it} = var(\mathbf{b}'_{it}\mathbf{F}_t) = b^2$ and again assuming that default probabilities depend only on the rating $r(i,t)$ we have

$$P(Y_{it} = 1 \mid F_t) = \Phi\left(\frac{\Phi^{-1}(p_{r(i,t)})}{\sqrt{1-b^2}} - \frac{bF_{ts(i,t)}}{\sqrt{1-b^2}}\right).$$

This model has GLMM structure

$$P\left(Y_{it} = 1 \mid \Psi_t\right) = \Phi\left(\gamma_{r(i,t)} + \Psi_{ts(i,t)}\right)$$

where $\boldsymbol{\Psi}_t = (\Psi_{t1}, \ldots, \Psi_{tp})' \sim N_p(\mathbf{0}, \Sigma)$ where Σ has diagonal elements $\sigma^2 + \tau^2$ and off-diagonal elements σ^2.

Equating the two models we get that

$$\begin{aligned}
\sigma^2 + \tau^2 &= \mathrm{var}(\Psi_{ts}) = \frac{b^2}{1-b^2} \\
\sigma^2 &= \mathrm{cov}(\Psi_{ts_i}, \Psi_{ts_j}) = \frac{\rho b^2}{1-b^2}
\end{aligned}$$

from which it follows that

$$\begin{aligned}
b^2 &= \frac{\sigma^2 + \tau^2}{1 + \sigma^2 + \tau^2} \\
\rho &= \frac{\sigma^2}{\sigma^2 + 1} \\
p_r &= \Phi\left(\frac{\gamma_r}{\sqrt{1 + \sigma^2 + \tau^2}}\right).
\end{aligned}$$

b^2 is the implied within-sector asset correlation, whereas ρ is the across-sector asset correlation.

2.2.4 Model 2.4: Model with sector random effects and macroeconomic covariates

Now consider the model with sector random effects is extended by adding observed macroeconomic covariates $\boldsymbol{z}_t$.

This model has GLMM structure

$$P\left(Y_{it}=1 \mid \boldsymbol{\Psi}_t\right)=\Phi\left(\gamma_{r(i,t)}+\boldsymbol{\eta}\boldsymbol{z}_t+\Psi_{ts(i,t)}\right)$$

where $\boldsymbol{\Psi}_t=(\Psi_{t1},\ldots,\Psi_{tp})'\sim N_p(\mathbf{0},\Sigma)$ where Σ has diagonal elements $\sigma^2+\tau^2$ and off-diagonal elements σ^2.

Equating the two models we get that

$$\begin{aligned}\sigma^2+\tau^2 &= \operatorname{var}(\Psi_{ts})=\frac{b^2}{1-b^2}\\ \sigma^2 &= \operatorname{cov}(\Psi_{ts_i},\Psi_{ts_j})=\frac{\rho b^2}{1-b^2}\end{aligned}$$

from which it follows that

$$\begin{aligned}b^2 &= \frac{\sigma^2+\tau^2}{1+\sigma^2+\tau^2}\\ \rho &= \frac{\sigma^2}{\sigma^2+1}\end{aligned}$$

Again we can also consider the correlation coming from macroeconomic covariates, b^2 is the implied within-sector asset correlation, whereas ρ is the across-sector asset correlation.

2.3 Empirical studies of default count data

2.3.1 Data description

A subset of the Standard & Poor's database CreditPro 6.6 which consists of 6897 US obligors from 13 industry sectors has been used in this analysis. The default count data have been collected for semester-based periods rating from January 1981 to December 2003 giving a total of $T=46$ periods. The yearly-based period miss many default events because many obligors migrate many times within one year, therefore we will misreport the default rating, while quarterly data improves the accuracy but too few default events in quarterly period. After comparing these three different data structure, we choose semester-based period in this default study with GLMMs. The empirical default probabilities are presented in Figure 2.1. All the rating-categories have many defaults at the same time periods and followed by other periods with few defaults.

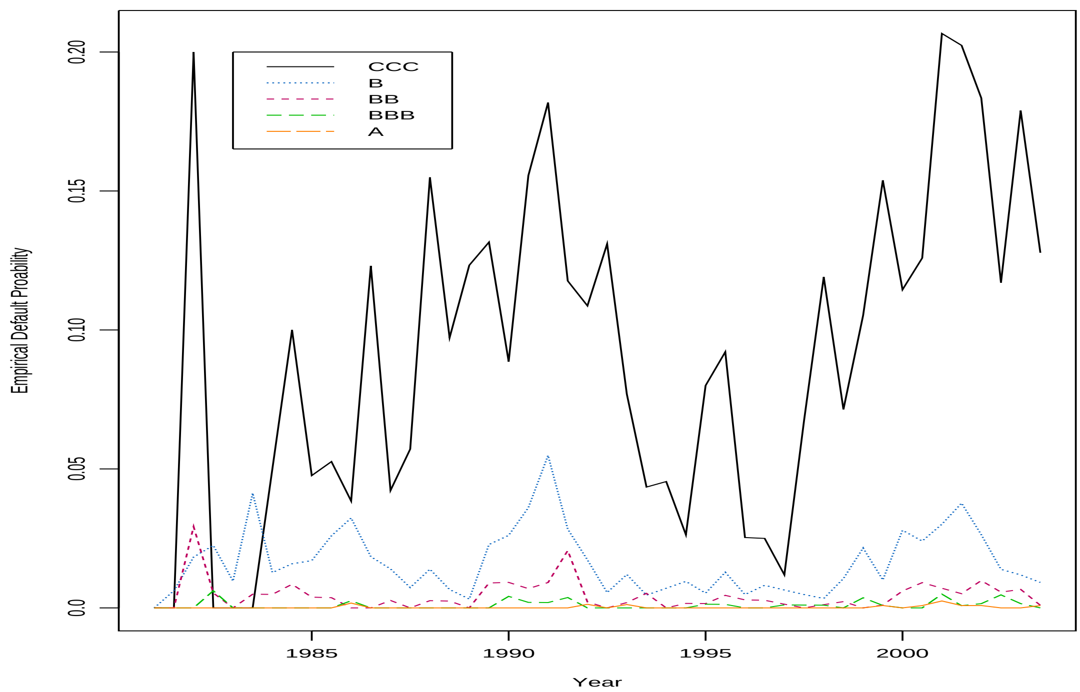

Figure 2.1: Empirical 6-month default rates for 23 years

The not-rated (NR) category in the S&P database is somewhat problematic. Some authors of empirical studies (e.g. Nickell et al. 2000) simply omit issuers who become NR from consideration. Other authors (e.g. Lando and Skoedeberg (2002)) treat transition to the not-rated category as a censoring event. Lando justifies this by citing evidence (Carty 1997) that the majority of transitions to not-rated are not related to changes in credit quality. In other words he argues that transition to not-rated is a non-informative censoring event, unrelated to the hazard of a firm subsequently defaulting. There is also the issue that many NR firms subsequently regain a rating or are recorded as defaulting at a later date. Although we are ignorant about their true rating during the time that they are NR, it seems a pity to have to exclude them from a default analysis in particular.

We conducted the following analysis of the NR phenomenon. We considered all transitions in the S&P database from ratings other than NR. We treated becoming NR as the event of interest and all other transitions (including default) as censoring events. We examined the covariates rating, country and sector. The conclusions: the lower rating categories have a significantly higher hazard of becoming NR than the higher rating categories; country has no discernable effect; tech firms have a significantly higher risk of becoming NR and utilities have significantly lower risk.

Thus the firms that become NR tend to be dominated by firms that have a higher default risk because they are lower rated. However, the question that is relevant from the point-of-view of the non-informative censoring assumption is whether a firm with a particular rating that becomes NR at a particular time point has a different default risk to a firm with the same rating that does not become NR? And there are many reasons to make a firm to become non-rated, including expiration of the debt, calling of the debt, failure to pay the requisite fee to S&P, etc. It is impossible for us to identify the exact reason for companies rating NR. Then we will treat NR case as censored in our analysis which means obligors whose rating is NR have been excluded from consideration but reconsider after they regain a rating.

The rating classes included in our analysis are

$$\mathcal{K} = \{CCC, B, BB, BBB, A\}$$

where we merge actual rating k^+, k, k^- into k. We also merge CCC, CC, and C into a single rating class CCC. Rating classes AAA and AA which rarely default have been excluded from this study; they will be reconsidered in our transition analysis in the following chapters. More than 75% of the companies in the dataset from US, so

Sector	Name	#Observations
1	"Aerospace/automotive/capital goods/metal"	837
2	"Consumer/service sector"+"Transportation"	1355
3	"Energy and natural resources"+"Utility"	994
4	"Financial Institutions"+"Insurance"+"Real estate"	1666
5	"Hightec/computers/office equipment"+"Telecom"	628
6	"Leisure time/media"	670

Table 2.1: Numbers of industry sectors observations

we choose a subset of the dataset which consists of 6150 US companies from the 6 selected industry sectors in our definition, see Table 2.1:

The industry sectors in our analysis are slightly different from the S&P definition of industry sectors. We merged Consumer/service with Transportation, Energy with Utility, Financial institution with Insurance and Real estate, and finally High technology with Telecommunications. The industry sectors we merged have broadly similar business and may be supposed to be similarly impacted by macroeconomic covariates.

2.3.2 Results

Without considering obligor-level covariates, counterparties are grouped according to rating category. We include an observed macroeconomic variable to partly explain the time-heterogeneity in the default rates. Several macroeconomic variables including Chicago Fed National Activity Index (CFNAI)are used. This helps us to detect lags between the cycle of index and the actual default cycle.

Model 2.1: One-factor model with Equicorrelation structure

Let $r = 1, \ldots, 5$ index the five rating categories in our study and γ_r is a fixed rating effect. The rating class default probabilities are given by $p_r = \Phi\left(\frac{\gamma_r}{\sqrt{1+\sigma^2}}\right)$. The random effect for period t is given by $\Psi_t \sim N(0, \sigma^2)$.

Model 2.1: GLMMs with Equicorrelation structure							
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	η	σ
mean	-3.5456	-3.1352	-2.7002	-2.1817	-1.2517	–	0.21629
s.e.	(0.1062)	(0.069)	(0.0518)	(0.0394)	(0.0461)	–	–
p_r	0.00026	0.00109	0.00416	0.01649	0.11059	–	–

Table 2.2: Model 2.1: GLMMs with equicorrelation structure fitted to 23 years historical Standard & Poor's data

The results for Model 2.1 are shown in Table 2.2. With this simplest GLMMs structure, each rating category has constant empirical default probability. However, the default risk might be affected by economic the cycle effects and so macroeconomic covariates will be added in the following analysis.

Model 2.2: Model with macroeconomic covariates

The numbers of defaults depend on the "state-of-the-economy". There are several possible proxies for "state-of-the-economy" which can possibly capture economic cycle effects. We investigate five different macroeconomic covariates in this analysis in order to find the best explanatory macroeconomic covariates. These covariates are growth in real GDP, return on the S&P 500 index (spretl), volatility of return on the S&P 500 index and Chicago Fed National Activities Index (CFNAI). Chava, Stefanescu and Turnbull (2008) use the S&P 500 index trailing one year return, GDP and CFNAI as explanatory variables. CFNAI has also been used in McNeil and Wendin (2007).

S&P 500 return (spretl in the following content) is the cumulated monthly return for the last 12 months at each time period t (the time period t represents semester). This is the same as in Chava, Stefanescu and Turnbull(2008). The greater the return in the previous year, the stronger the economy and the lower the probability of default, so that a negative value is expected for the macroeconomic effect. Default probabilities

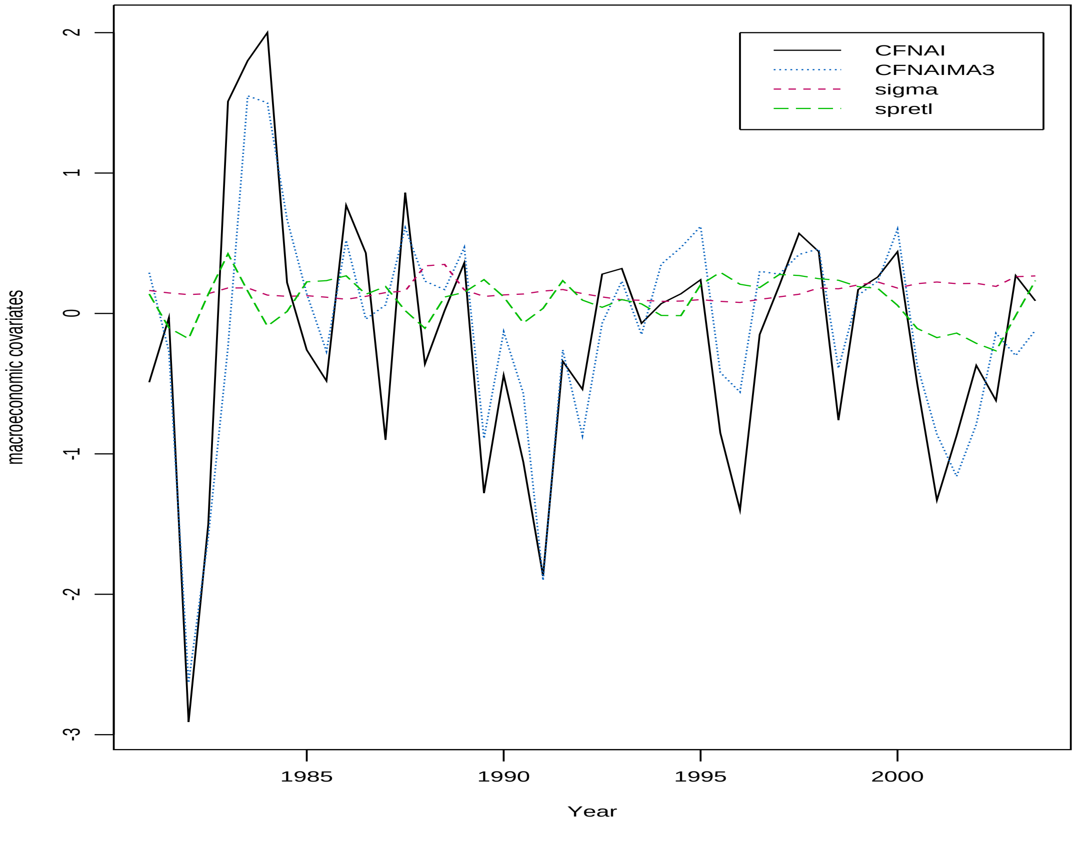

Figure 2.2: Historical macroeconomic covariates CFNAI, CFNAIMA3, SP500 return and SP500 volatility

should decrease with increasing S&P 500 return. The growth in real GDP (ΔGDP) also has an expected negative effect. The greater the growth in the economy means most companies would make more money and have a lower probability of default. The ΔGDP we use first quarter of each time period t. The higher volatility means higher risk, so we expected a positive relationship for volatility of S&P return and default probability. Chicago Fed National Activity Index (CFNAI), is a weighted average of 85 existing, monthly indicators of national economic activity. The CFNAI provides a single, summary measure of a common factor in these national economic data. Historical movements closely track periods of economic expansion and contraction. This monthly released index tracks periods of increasing and decreasing inflationary pressures. The CFNAIMA3 is the three months moving average of CFNAI which tracks economic expansions and contractions. The CFNAI is a coincident indicator of economic expansions and contractions. Default events normally happen later than the economic downturn. After a simple empirical on CFNAI, its two and three months moving average, we find three month moving average fits the model better than the other two. We will focus on the CFNAIMA3 rather than CFNAI. As with index return and GDP, a negative relationship is expected here.

All these five macroeconomic covariates coefficients gave the expected sign, see Table 2.3. We get negative coefficient for ΔGDP, spretl, CFNAI and CFNAIMA3 and positive coefficient for stock index return volatility. Probabilities of default are given by $P\left(Y_{it}=1 \mid \Psi_t\right)=\Phi\left(\gamma_{r(i,t)}+\boldsymbol{\eta}\boldsymbol{z}_t+\Psi_t\right)$. The negative coefficient will decrease the default probabilities with increasing observed macroeconomic covariates while positive coefficient will increase the default probabilities.

All of ΔGDP, spretl, CFNAI and CFNAIMA3 have statistically significant effects while volatility has not. Among them, the p-value for spretl, CFNAI and CFNAIMA3 are smaller than .005 which means statistically significant for 99.5% confidence interval. The p-value for volatility is 0.14 which is not statistically significant. The CFNAI index is a weighted average of 85 existing, monthly indicators of US national economy activity which can explain the US economic. S&P is used in reference not only to the

Model 2.2: GLMM with macroeconomic covariate ΔGDP							
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	η	σ
mean	-3.4361	-3.0253	-2.5899	-2.0706	-1.1414	-0.0323	0.1960
s.e.	(0.1127)	(0.0787)	(0.0642)	(0.0551)	(0.0601)	(0.0123)	–
p-value	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	0.0119	–

Model 2.2: GLMM with macroeconomic covariate spretl							
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	η	σ
mean	-3.4873	-3.0777	–2.6426	-2.1231	-1.1947	-0.6293	0.1898
$s.e.(\gamma_r)$	(0.1071)	(0.0698)	(0.0526)	(0.0405)	(0.0471)	(0.2120)	–
p-value	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	0.0048	–

Model 2.2: GLMM with macroeconomic covariate volatility							
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	η	σ
mean	-3.6875	-3.2777	-2.8424	-2.3242	-1.3950	0.8801	0.2115
s.e.	(0.1427)	(0.1179)	(0.1085)	(0.1031)	(0.1060)	(0.5858)	–
p-value	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	0.1400	–

Model 2.2: GLMM with macroeconomic covariate CFNAI							
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	η	σ
mean	-3.5682	-3.1575	-2.7229	-2.2033	-1.2735	-0.1186	0.1909
s.e.	(0.1064)	(0.0686)	(0.0507)	(0.0376)	(0.0446)	(0.0400)	–
p-value	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	0.0049	–

Model 2.2: GLMM with macroeconomic covariate CFNAIMA3							
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	η	σ
mean	-3.5646	-3.1538	-2.7185	-2.1982	-1.2694	-0.1765	0.1681
s.e.	(0.1063)	(0.0675)	(0.0488)	(0.0346)	(0.0423)	(0.0426)	–
p-value	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	0.0001	–

Table 2.3: Model 2.2: GLMM with different macroeconomic covariates

index but also to the 500 companies that have their common stock included in the index and S&P default data used in our analysis. That is why these three observed macroeconomic covariates have the highest explanatory power. See Figure 2.2 for historical macroeconomic covariates. We will compare these three observed factor with unobserved factor as well.

We compared the observed factor in Model 2.2 with the unobserved factor in Model 2.2 and the two unobserved factors in both Model 2.1 and Model 2.2. The upper plot of Figures 2.3, 2.4 and 2.5 displays the unobserved factor $\{(t, \Psi_t) : t = 1, \ldots, T\}$ of Model 2.1 (solid black line) and observed fixed factors $\{(t, z_t\eta) : t = 1, \ldots, T\}$(dashed blue line) in Model 2.2. The lower plot compare the unobserved factor for Model 2.1 (solid black line) and Model 2.2 with macroeconomic covariates(dashed blue line). There seems to be a fair amount of co-movement between the two series $\{(t, z_t\eta) : t = 1, \ldots, T\}$ and $\{(t, \Psi_t) : t = 1, \ldots, T\}$, but it is obvious that z_t does not track Ψ_t particularly accurately, and that z_t does not fully capture the default activity. This illustrates the problems associated with observed proxies for the systematic risk. The lower plots compare unobserved factor of Models 2.1 and 2.2. Although the paths are very similar, the reduced variance of Ψ_t when z_t is explicitly modelled can be detected. the standard deviation is reduced from 0.2163 in Model 2.1 to 0.1681 (CFNAIMA3), 0.1898 (spretl) and 0.1909 (CFNAI) in Model 2.2. This can be easily detected from the plots, the random effects for spretl varies more than CFNAI and CFNAIMA3. These results help to understand that CFNAIMA3 is the best single macroeconomic covariate of the ones we have tried to explain US economy.

We studied all these five macroeconomic covariates individually. In order to get further comparison, we analysis all these macroeconomic covariates in one model. Volatility is removed from our consideration because it is not statistically significant in previous analysis. CFNAIMA3 is three months moving average of CFNAI, so we pick CFNAIMA3 instead of CFNAI. See the results in Table 2.4.

From Table 2.4, both CFNAIMA3 and spretl have statistically significant effects while GDP has not. Both CFNAIMA3 and spretl give a negative coefficient which is ex-

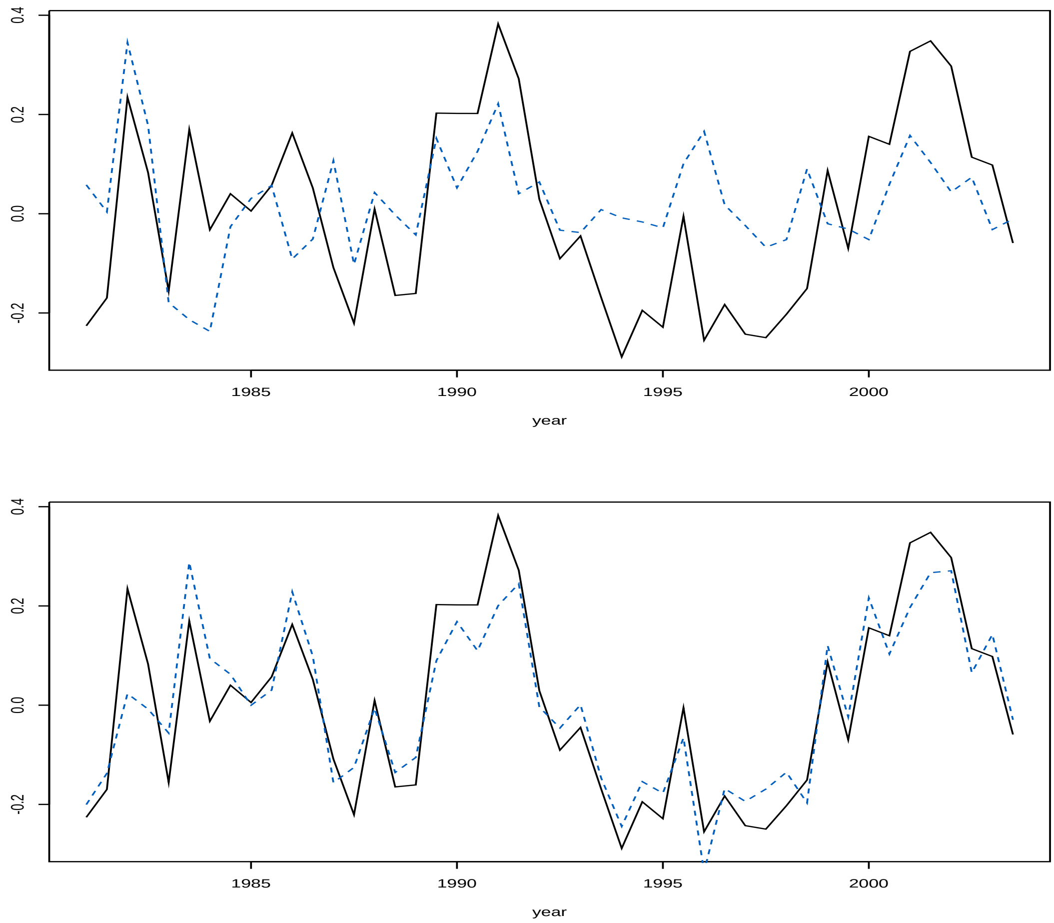

Figure 2.3: Visual comparison of systematic risk factors. The upper plot displays the estimated unobserved effect of factor $\{(t, \Psi_t) : t = 1, \ldots, T\}$ in Model 2.1 (solid black line) and the estimated fixed factor CFNAI $\{(t, z_t\eta) : t = 1, \ldots, T\}$(dashed blue line) in Model 2.2. The lower plot compare the unobserved factor for Model 2.1 (solid black line) and Model 2.2 with macroeconomic covariates CFNAI (dashed blue line).

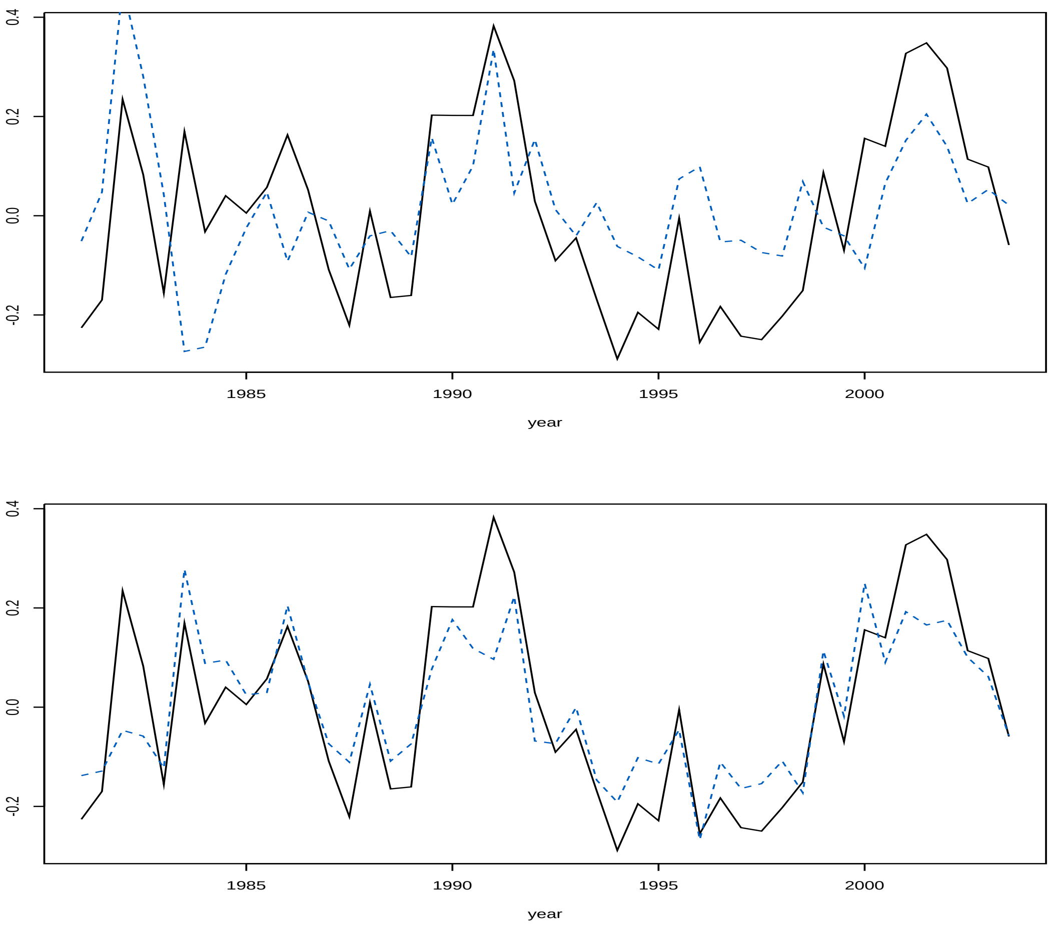

Figure 2.4: Visual comparison of systematic risk factors. The upper plot displays the unobserved factor $\{(t, \Psi_t) : t = 1, \ldots, T\}$ of Model I (solid black line) and observed fixed factor CFNAIMA3 $\{(t, z_t\eta) : t = 1, \ldots, T\}$(dashed blue line) in Model 2.2. The lower plot compare the unobserved factor for Model I (solid black line) and Model 2.2 with macroeconomic covariates CFNAIMA3(dashed blue line).

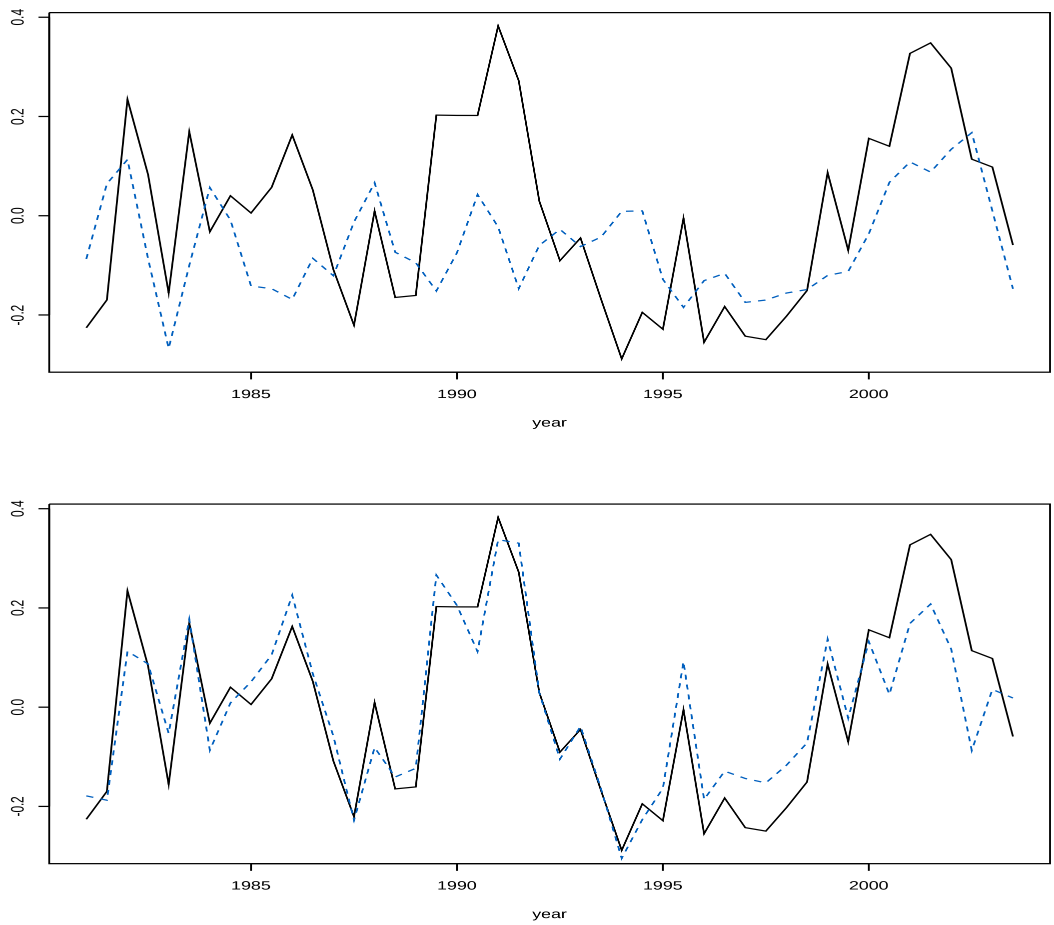

Figure 2.5: Visual comparison of systematic risk factors. The upper plot displays the unobserved factor $\{(t, \Psi_t) : t = 1, \ldots, T\}$ of Model 2.1 (solid black line) and observed fixed factor Spretl $\{(t, z_t\eta) : t = 1, \ldots, T\}$(dashed blue line) in Model 2.2. The lower plot compare the unobserved factor for Model 2.1 (solid black line) and Model 2.2 with macroeconomic covariates spretl(dashed blue line).

pected. However, GDP gives a positive relationship which is different from our expectation. GDP worked well in our previous analysis but not this model. This result tells us that CFNAIMA3 and spretl explain the US economy better than GDP. CFNAIMA3 explained US economy better than spretl. Both CFNAIMA3 and spretl are statistically significant in our results, although p-value for CFNAIMAS is smaller than spretl. Here we fit the model for only both two macroeconomic covariates.

From Table 2.5, both CFNAIMA3 and spretl have statistically significant effects. Both CFNAIMA3 and spretl give a negative coefficient which is expected. This result tells us that CFNAIMA3 and spretl explain the US economy better than GDP. CFNAIMA3 explained US economy better than spretl with a smaller p-value for CFNAIMAS. Macroeconomic covariates spretl did some additional explanation, but CFNAIMA3 is the most important covairtes to explain the credit quality changes, we will use the CFINAMA3 as the observed factor in further analysis.

Model 2.3: Model with sector random effects

GLMMs allow additional random effect to capture patterns of variability in the response which cannot be explained by observed covariates. Random effects like industry sector can be added in order to capture additional variability. We use 6 industry sectors out of 8 industry sector which we created using CreditPro data.

With the additional random effect, the maximum value of the log-likelihood is -1818 in the model with industry sectors and -1913.3 in the basic model, see Table 2.8. The systematic risk can be thought of as being divided into two parts which are explained by fixed effects and random effects. With the additional sector random effect, heterogeneity is allowed in default rates between sectors within a time period. Further random effects like country can be introduced to allow more heterogeneity, although we prefer to concentrate on US data.

The default probabilities in models with sector random effect are given by $p_r = \Phi\left(\frac{\gamma_r}{\sqrt{1+\sigma^2+\tau^2}}\right)$.[1] We compare with model with sector random effects and the model

[1] The Equicorrelation Structure here is the same as before. However, we split the data into 6

Model 2.2 with macroeconomic covariate CFNAIMA3 spretl and GDP						
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	η_1
mean	-3.5406	-3.1306	-2.6952	-2.1739	-1.2465	-0.1643
s.e.	(0.1275)	(0.0970)	(0.0849)	(0.0775)	(0.0810)	(0.0686)
p-value	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	0.0210
Remaining Parameters						
	η_2	η_3				
mean	-0.4319	0.0049				
s.e.	(0.1939)	(0.0175)				
p-value	0.0312	0.7809				

Table 2.4: Model 2.2 with macroeconomic covariate CFNAIMA3 spretl and GDP

Model 2.2 with macroeconomic covariate CFNAIMA3 and spretl							
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	η_1	η_2
mean	-3.552	-3.095	-2.683	-2.138	-1.205	-0.155	-0.454
s.e.	(0.1044)	(0.0643)	(0.0497)	(0.0374)	(0.0443)	(0.043)	(0.190)
p-value	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	0.0008	0.0215

Table 2.5: Model 2.2 with macroeconomic covariate CFNAIMA3 and spretl

Model 2.3: GLMMs with industry sectors							
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	σ	τ
mean	-3.5806	-3.1774	-2.7606	-2.2445	-1.2834	0.1945	0.2571
s.e.	(0.0936)	(0.0626)	(0.0484)	(0.0391)	(0.0437)	–	–
p-value	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	–	–
p_r	0.00033	0.00125	0.00430	0.01633	0.11094	–	–

Table 2.6: Model 2.3 with industry sectors random effects

without sector-specific random effects. Model with sector random effects has lower *Akaike Information Criterion (AIC)*$(AIC = -2 * logL + 2 * k)$ and *Bayesian Information Criterion (BIC)* $(AIC = -2 * logL + k * logN)$ than model without-specific random effects, see Table 2.8.

The within-sector implied asset correlation is 9.41%, whereas the across-sector counterpart is only 3.42%. The model without sector-specific random effects has an overall implied asset correlation 4.47%. Implied asset correlation is obviously increased within an industry sector.

The estimated industry sector random effects show both similarities and difference. From Figure 2.6 and 2.7, both the industry sectors and empirical default probabilities have the same peak around 1990. However, around 2000, the "Hightec" industry has another peak while "Energy" and "Finance" industry do not. The end of the high-technology speculative bubble starting at 2000 caused a lot of high technology companies to default. Industry "Aero","Consumer" also have peak while the end of bubble. Industry sector "Energy" has lower random effect around 1995 while other sectors have higher random effect. This industry sector-specific random effects capture the information which the simpler one-factor model could not and show difference for different industry sectors.

Model 2.4: Model with sector random effects and macroeconomic covariates

In the previous research, we find CFNAIMA3 is the best explanatory macroeconomic covariate for the US economy. Thus CFNAIMA3 will be used as a macroeconomic coavriate in this analysis. Read Table 2.9 for results.

The macroeconomic variable still has a statistically significance effect. The within-sector implied asset correlation is 8.03%, whereas he across-sector counterpart is only 2.19%. The model without sector-specific random effects has an overall implied asset

industry which make the data size is 6 time bigger than the data used before

Model 2.1: GLMMs with Equicorrelation structure						
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	σ
mean	-3.5439	-3.1336	-2.6988	-2.1806	-1.2507	0.2143
s.e.	(0.10890)	(0.07052)	(0.05251)	(0.03952)	(0.04660)	–
p-value	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	–
p_r	0.00026	0.00109	0.00416	0.01650	0.11068	–

Table 2.7: Model 2.1: GLMMs with Equicorrelation structure

	AIC	BIC	logLik
With industry sector random effect	3652.0	3693.8	-1818.0
Without industry sector random effect	3840.6	3877.2	-1913.3

Table 2.8: Model with sector random effects and without sector random effects

Model 2.4: GLMMs with industry sectors and CFNAIMA3						
	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	η
γ_r	-3.5959	-3.1929	-2.7760	-2.2589	-1.2992	-0.1616
$s.e.(\gamma_r)$	(0.09241)	(0.06045)	(0.04531)	(0.03489)	(0.04008)	(0.04347)
p-value	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	0.0006
Remaining Parameters						
	σ	τ				
γ_r	0.1497	0.2549				
$s.e.(\gamma_r)$	–	–				
p-value	–	–				

Table 2.9: Model 2.4 with industry sectors random effects and CFNAIMA3

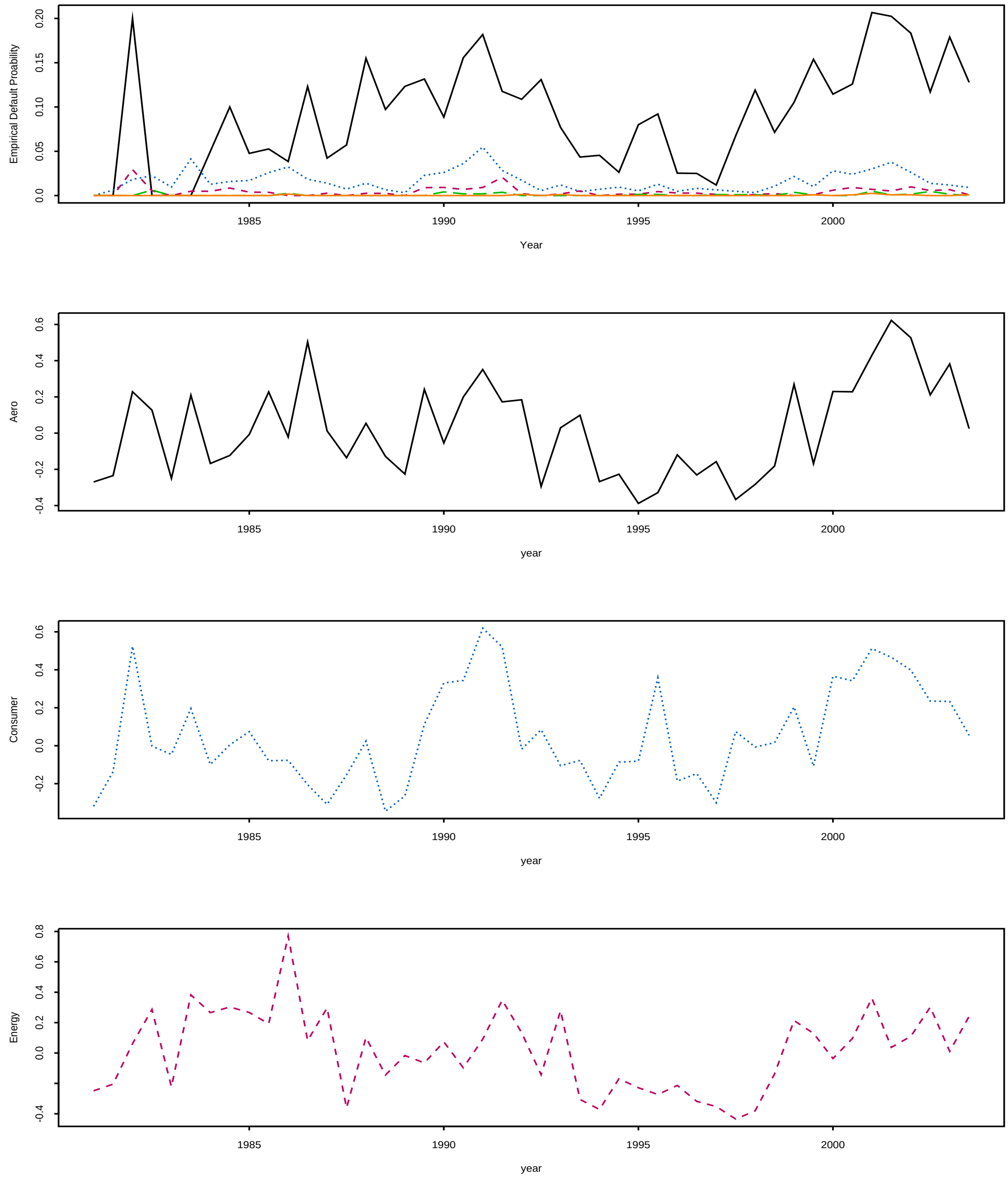

Figure 2.6: The top figure shows empirical default probability with different rating categories, the following three plots show $\{(t, b_t) : t = 1, \ldots, T\}$ for different industry Aero, Consumer and Energy random effects with rating category "CCC"

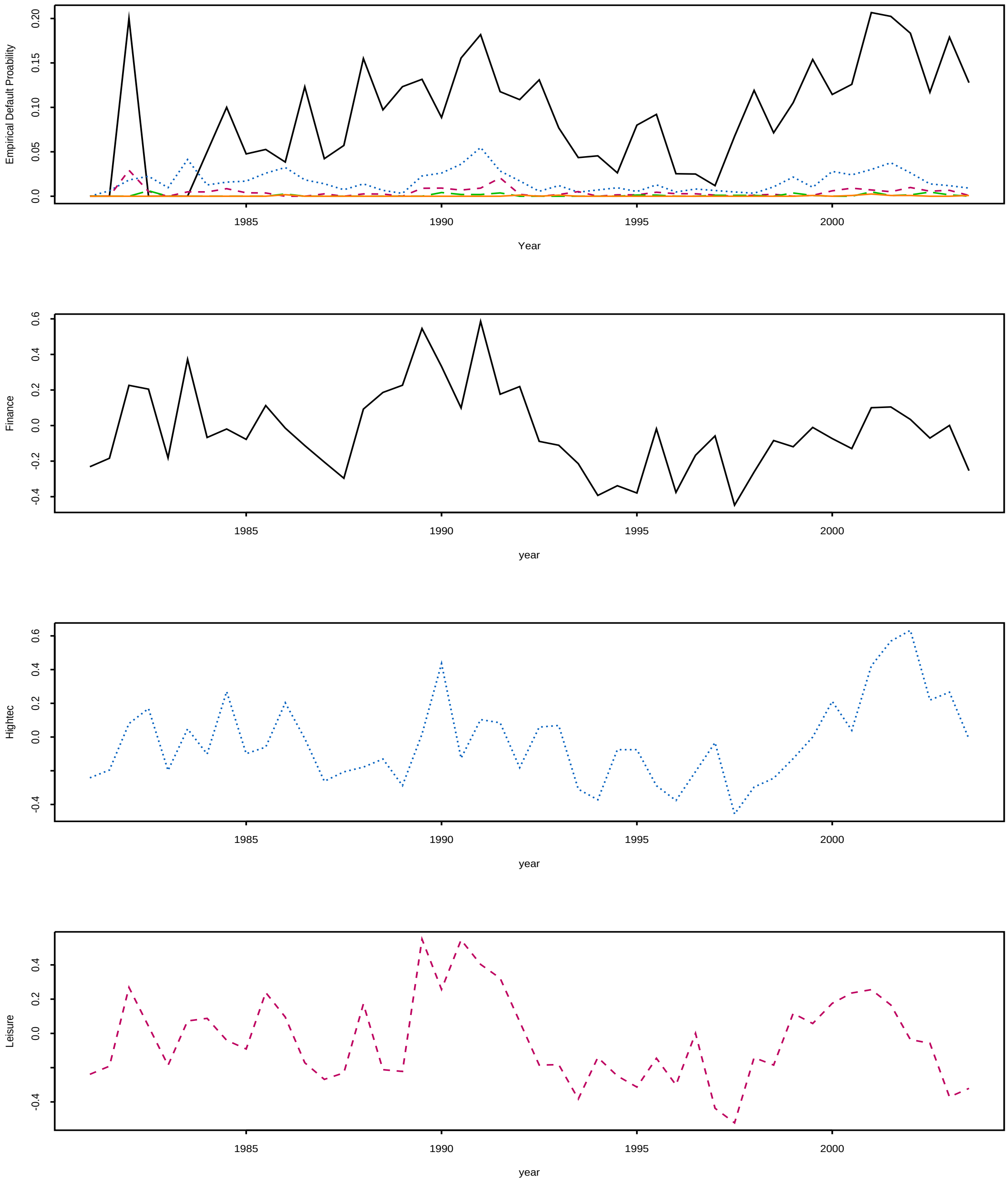

Figure 2.7: The top figure shows empirical default probability with different rating categories, the following three plots show $\{(t, b_t) : t = 1, \ldots, T\}$ for different industry Finance, Hightec and Leisure random effects with rating category "CCC"

correlation of 4.47%. Implied asset correlation increased within an industry sector. The standard deviation is 0.14968 which is reduced from 0.19446 in Model 2.3.

	AIC	BIC	logLik
Model 2.4	3653.6	3700.6	-1817.8
Model 2.3	3652.0	3693.8	-1818.0
Model 2.2	3867.2	3909.0	-1925.6
Model 2.1	3840.6	3877.2	-1913.3

Table 2.10: LogLik,AIC and BIC for four different models

Table 2.10 gives the logLik, AIC and BIC for three different models. The maximum log-likelihood are similar for model with and without macroeconomic covariate effect but massively increase with the industry sectors random effects. See Table2.10 for details. The industry sectors random effect increase the maximum log-likelihood from -1919.3 and -1925.6 to -1818.0 and -1817.8. But with additional observed macroeconomic variable in Model 2.3, the log-likelihood only increased 0.2 to -1817.8, the log-likelihood even decreased in Model 2.1 with additional observed macroeconomic variable. This result shows that industry sector random effect can help us to capture pattern of variability in response that cannot be explained by observed macroeconomic variables alone. The random industry effects even take more contribution than macroeconomic covariate for default acativity.

2.4 Discussion

GLMMs allow the systematic portfolio risk to be divided into observed fixed effects and unobserved random effects to capture heterogeneity in default rates. Multivariate random effects can even capture heterogeneity in time and across industry sectors or country.

We choose to use standard package S-plus in this analysis. The standard package can fit one factor credit risk model with equicorrelation structure. Further observed fixed effects like macroeconomic covariates and unobserved random effects can be included.

In our empirical study of the Standard and Poor's default data find CFNAIMA3 is the best observed macroeconomic covariates to describe the credit default among the macroeconomic covariates. Other observed covariates like spretl and CFNAI work well too. The variance of random effects has been reduced with the macroeconomic variables and thus the implied asset correlations. However, the macroeconomic variables cannot capture the full variability in default rates.

The empirical analysis also considers the industry-specific latent factors. Our results show increased implied asset correlations within industry sectors. This give us an important message that the issue of heterogeneity between industry sectors or country sectors should be considered in credit risk models.

However, the constructed credit portfolio with equicorrelation correlation structure will give the same loss distribution although the model has been proved to have improvement. Furthermore, the standard statistical package does not allow serial dependence. We need to be able to estimate more complicated models for our further research.

Modelling time-to-event with frailties (called random effects in GLMMs framework) rather than just numbers of events is one of the possibilities. We will give more details about the survival models for credit risk modelling in the following two chapters. Furthermore, we need to extend our default data to transition data. McNeil and Wendin (2007) study both default and transition probabilities, we will model default and transition risk with time-to-event models.

Chapter 3

Modelling default risk with survival models

For credit risk modelling, we are not only interested in the number of companies that migrate from one rating category to another, but also interested in the time period the company spends in a certain rating category. Hence, we will study the time-to-event analysis which has become popular in credit risk modelling in recent years. An analysis with GLMM models for credit risk modelling has been provided in the previous chapter. We will cover time-to-event methods in this chapter.

In his seminal work, Cox (1972) proposes the proportional hazards model, where it is possible to estimate the relative intensity of a decrement without specifying the baseline intensity. Cox (1975) demonstrates the estimation procedure for the proportional hazards model with partial likelihood estimation. Andersen and Gill (1982) generalize the model to allow time-varying covariates using a counting process formulation, and show that the maximum partial likelihood estimates are asymptotically equivalent to unconditional maximum likelihood estimates.

As in our count data analysis for GLMMs, frailty in survival models help to capture heterogeneity with unobserved random variables. We introduce a frailty-based survival model for modelling the intensity of credit rating transitions (default). This type of model is an extension of the Cox proportional hazards models where a common random variable is used to account for heterogeneity. Kavvathas (2001) and Couderc and Renault (2005) use a similar duration approach conditional on observed macro-variables and they show that average time-to-default decreases as economic activity decreases. Shumway (2001) develops a more dynamic bankruptcy prediction model by combining both financial ratios and market-driven measures and argues that discrete-time is necessary to calibrate hazards because of the intermittency of accounts information. Chava and Jarrow (2004) extend Shumway's (2001) analysis to consider industry sector heterogeneity using monthly intervals. Duffie, Saita, and Wang (2007) formulate a doubly stochastic model for firm survival using firm-specific and macroeconomic covariates.

We adapt a simpler model used in medical statistics (Manda and Mayer (2005)) and extended for the credit risk application. We estimate rating transition model with shared dynamic frailties for different industry sectors and macroeconomic covariates using Bayesian techniques (MCMC). This is a model that each transition intensity follows a Cox type multiplicative regression model with two levels of frailties to account for both time period and industry sector heterogeneity. Delloye, Fermanian and Sbai (2006) define a reduced-form credit portfolio model which treat rating transition as independent competing risks with conditionally independent and proportional hazards assumption. They also allow strong dependence levels by adding heterogeneity. However, Delloye, Fermanian and Sbai (2006) split the Standard & Poors' data into several groups based on the similar rating transition type and analyze these separately. We consider the whole transition data and let all rating transitions share the same macroeconomic covariates and unobservable random process for all companies in monthly interval.

We have two aims, First, we will compare the time-to-event model with GLMMs which were described in the previous chapter. Second, we will investigate differences between sectors. Cox's hazard model, which has been increasingly used to model the hazard of credit risk events in recent years, is used in this analysis. Cox (1972) proposed the proportional hazards model and Andersen and Gill (1982) generalized the model to allow time-varying covariates using counting process formulation.

In this chapter, we will briefly introduce the Cox model and Cox Proportional hazard model and their estimation methods. Then we will apply Manda and Meyer (2005) with time-to-event credit risk models. We will present application to Standard & Poor's CreditPro data with time-dependent frailty models for recurrent failure time data in Bayesian context and estimate it using the Markov Chain Monte Carlo methods. Bayesian methods based on Markov Chain Monte Carlo (MCMC) techniques have three advantages: Bayesian methods improve estimation accuracy, Bayesian estimation also allows for taking into account expert opinion through the use of subjective prior distribution of for model parameters, and Bayesian inference becomes straightforward to compute the default and transition probabilities. We will provide comparative results with credit risk models using GLMMs. The random effect in GLMMs models will appear as frailties in survival models. An autoregressive process is used as time-dependent frailty and two levels frailty time and industry will be used in this model. We will investigate differences in default intensity between industry sectors.

3.1 Theory

The Cox regression model for censored survival data specifies that covariates have a proportional effect on the hazard function of the life-time distribution of an individual. Anderson and Gill (1982) extend the Cox model to a model where covariate processes have a proportional effect on the intensity process of a multivariate counting process. This allows complicated censoring patterns and time dependent covariates.

3.1.1 Models

Cox proportional hazard model

Cox (1972) specifies the hazard rate or intensity of failure $\lambda_i(t) = \lim_{h\downarrow 0} P[T_i \leq t+h \mid T_i > t]$ for the survival time T_i of an individual with covariate vector $\boldsymbol{z}_i$ to have the form

$$\lambda_i(t) = \lambda_i(t|\boldsymbol{\beta}) = \lambda_0(t)\exp(\boldsymbol{\beta}'\boldsymbol{z}_i), \quad t \geq 0. \tag{3.1}$$

here $\boldsymbol{\beta}$ is a $p-$vector of unknown regression coefficients and $\lambda_0(t)$ is the underlying hazard which is an unknown and unspecified nonnegative function.

Andersen-Gill model

Andersen and Gill (1982) discuss how the Cox model can be extended to a model where covariate processes have a proportional effect on the intensity process of a multivariate counting process. The A-G model permits a statistical regression analysis of the intensity of a recurrent event and allow for complicated censoring patterns and time dependent covariates. This is relevant for credit risk modelling.

Consider any subject i is followed over time and can experience multiple events of the same type such that the times of events are ordered $0 < T_{i1} < T_{i2} < \cdots$, with the probability 0 of tied observations. If the event is default, it is of course unlikely that we would have multiple events but the framework allows this possibility. The predetermined time interval $[0, c_i]$ is divided into discrete time interval. In practice monthly intervals prove to be sufficiently small. At time t, we have for the i^{th} subject a d-dimensional vector of risk factors $\boldsymbol{X}_i(t)$, and an observable process $N_i(t)$ that counts the number of events which have occurred up to time t; and $Y_i(t)$, a non-negative predictable indicator process taking the value 1 if the subject is under observation and 0 otherwise. In modelling such counting process with observed data $D = (N, X, Y)$, we look at the intensity of the process which measures the risk of an event at time t.

This intensity is modelled as time dependent. For the i^{th} subject, the corresponding intensity at time t is $\lambda_i(t|\boldsymbol{\beta})$, where $\boldsymbol{\beta}$ is a vector of unknown parameters which are associated with the covariate vector $\boldsymbol{X}_i(t)$. This model can be extended to a frailty-based model which can be used for modelling intensity of credit rating transitions. It uses a common random variable to account for the heterogeneity in intensity rates that can not be accounted for by observed covariates.

Andersen-Gill model with static frailty

The frailty-based Andersen-Gill model will be extended by adding a subject-specific random frailty W_i which can capture the risk not accounted for by the included observed risk variables $\boldsymbol{X}_i(t)$. Then the intensity at time t will be $\lambda_i(t|\boldsymbol{\beta}, W_i)$.

The study period for subject i is partitioned into finite disjoint intervals $A_t = [t, t + dt), 0 \leq t \leq c_i$ such that no time interval contains more than one of the consecutive events of the subject, then the counting process jump $dN_i(t) = N_i(t+dt) - N(t)$ takes only value 0 or 1. The intensity function is given by

$$\lambda_i(t|\boldsymbol{\beta}, W_i)dt = P(dN_i(t) = 1|F_{t-}; \boldsymbol{\beta}, W_i) \tag{3.2}$$

where F_{t-} is the available data just before time t. Andersen and Gill (1982) give the intensity function

$$\lambda_i(t|\boldsymbol{\beta}, W_i) = Y_i(t)\lambda_0(t)\exp(\boldsymbol{\beta}'\boldsymbol{X}_i(t) + W_i) \tag{3.3}$$

where $Y_i(t)$ is indicator function, $\lambda_0(t)$ is an baseline intensity function and $\exp(\boldsymbol{\beta}'\boldsymbol{X}_i(t)+ W_i)$ is the Cox covariate effect function. The advantage of this intensity model is that it allows time-varying covariates by risk factor $\boldsymbol{X}_i(t)$ and also include subject-specific frailty W_i on the baseline hazard.

Andersen-Gill model with dynamic frailty

In frailty models, an independent subject-specific random effect is usually assumed to be time-constant for each subject. Manda and Meyer (2005) present application to medical data with time-dependent frailty model for recurrent failure time data in the Bayesian context and estimate it using the Markov Chain Monte Carlo method; an autoregressive process is used as time-dependent frailty. The intensity function with time-dependent frailty is extended as

$$\lambda_i(t|\boldsymbol{\beta}, W_i(t)) = Y_i(t)\lambda_0(t)\exp(\boldsymbol{\beta}'\boldsymbol{X}_i(t) + W_i(t)) \tag{3.4}$$

where $W_i(t)$ time-dependent subject specific frailty. In our credit risk application, we use dynamic shared group frailties instead of subject-specific frailties.

3.1.2 Parameter estimation

Cox originally proposed a *partial likelihood* approach. There are several methods for estimating the Andersen-Gill formulation of the Cox proportional hazard model. Like Maximum likelihood estimation which is available in standard R package, however, these function only can estimate some simple models. Clayton (1991) formulates the Cox model using the counting process notation introduced by Andersen and Gill (1982) and discusses estimation of the baseline hazard and regression parameters using MCMC methods. Manda and Meyer (2005) also use the Bayesian inference for recurrent events using time-dependent frailty.

In this thesis, we use Gibbs simulation which is probably the most widely used M-CMC methods for estimating our different models. The BUGS (Bayesian inference Using Gibbs Sampling) is concerned with flexible software for the Bayesian analysis of complex statistical models using Markov chain Monte Carlo (MCMC) methods. WinBUGS is part of the BUGS project, WinBUGS can use either a standard 'point-and-click' windows interface for controlling the analysis, or can construct the model using a graphical interface called DoodleBUGS. R2WinBUGS is a package running

WinBUGS from R. Using this package, it is possible to call a BUGS model, summarize inferences and convergence in a table and graph, and save the simulations in arrays for easy access in R.

Partial likelihood

Cox (1975) originally proposed *partial likelihood* for proportional hazards model estimation. Although it is now well known that partial likelihood is misnamed and not a likelihood but seemingly based on standard likelihood results.

Suppose we have n possibly right censored survival times $T_1, \cdots, T_n$ and the corresponding covariate vectors $\boldsymbol{z_1}, \cdots, \boldsymbol{z_n}$, where $\boldsymbol{z_i}$ is observed on $[0, T_i]$. Cox (1972) suggested that inference on $\boldsymbol{\beta}$ be based on the function

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{n} \left\{ \frac{e^{\boldsymbol{\beta}' \boldsymbol{z}_i}}{\sum_{j \in \mathcal{R}_l} e^{\boldsymbol{\beta}' \boldsymbol{z}_j}} \right\}^{\delta_i} \tag{3.5}$$

where $\mathcal{R}_l = \{j : T_j \geq T_i\}$ is the risk set and δ_i is an indicator for failure and $1-\delta_i$ is an indicator for censoring. Cox (1975) derived the above formula as a partial likelihood function. Letting $\hat{\boldsymbol{\beta}}$ be the value that maximizes (3.5), then the continuous estimator obtained by linear interpolation between failure times of

$$\hat{\Lambda}(t) = \sum_{T_i \leq t} \frac{\delta_i}{\sum_{j \in \mathcal{R}_l} e^{\hat{\boldsymbol{\beta}}' \boldsymbol{z}_j}} \tag{3.6}$$

for the underlying cumulative hazard $\Lambda_0(t) = \int_0^t \lambda_0(s) ds$ was suggested by Breslow(1972,1974).

Full maximum likelihood procedure

We consider default risk only in this chapter, but will extend the model to the migration case in the next chapter. Here we will only consider the full maximum likelihood procedure for default model and will show in the next chapter for migration case. In general, let $\lambda_i(t)$ be the intensity for firm i . In formula (3.1), λ_0 is the baseline hazard function, $\boldsymbol{\beta}$ is the parameter to be estimated and $\boldsymbol{z}$ is the macroecnomic covariates in this analysis. In our simple model, covariate only depend on time t. The observable process $N_i(t)$ denote the number of event for firm i at period time t. $Y_i(t)$ be the indicator variable which take value 1 or 0 depend on under observation or not.

Andersen et al. (1993) proposed a full maximum likelihood procedure for estimating the unknown coefficients. The details can be found in Andersen et al. (1993), Kavvathas (2000) or Delloye et al (2005). With n firms, the likelihood could be written as following:

$$\mathcal{L} = \prod_{i=1}^{n} \mathcal{L}_i$$

$$\mathcal{L}_i = \left\{ \prod_t (\lambda_i(t|\boldsymbol{z}))^{dN_i(t)} \right\} \cdot \left(- \int_0^\infty Y(u)\lambda_i(u|z)du \right) \tag{3.7}$$

The log-likelihood will be:

$$\ln \mathcal{L}_i = \sum_{i=1}^{n} \ln \mathcal{L}_i$$

Bayesian inference

Several authors have discussed Bayesian inference for frailty models. Clayton(1991) set out a Bayesian representation of the model and discussed inference using Monte Carlo methods. We propose Manda and Meyer (2005) model in our analysis, thus the Bayesian inference using time-dependent frailty will be discussed. The joint distribution of the observed data D given $\lambda_0, \boldsymbol{\beta}, W$ is given by

$$\begin{aligned} p(D|\lambda_0, \boldsymbol{\beta}, W) &= \prod_i^I \left\{ \prod_{t\geq 0}^{c_i} [Y_i(t)e^{\boldsymbol{\beta}' \boldsymbol{X}_i(t)+W_i(t)}\lambda_0(t)]^{dN_{i(t)}} \right\} \\ &\quad \times \exp\left(- \int_0^{c_i} Y_i(t)e^{\boldsymbol{\beta}' \boldsymbol{X}_i(t)+W_i(t)}\lambda_0(t)dt \right) \\ &\propto \prod_i^I \left\{ \prod_{t\geq 0}^{c_i} [e^{\boldsymbol{\beta}' \boldsymbol{X}_i(t)+W_i(t)} d\Lambda_0(t)]^{dN_{i(t)}} \exp(-Y_i(t)e^{\boldsymbol{\beta}' \boldsymbol{X}_i(t)+W_i(t)} d\Lambda_0(t)) \right\} \end{aligned}$$

As detailed in Manda and Meyer(2005), the purpose of Bayesian analysis for frailty is to determine the summary statistics of the posterior distribution of parameters $\boldsymbol{\psi}$. The posterior distribution is from $p(\boldsymbol{\psi}|D) \propto p(D|\boldsymbol{\psi})p(\boldsymbol{\psi})$ and the data likelihood above is updated with the prior distribution of the model parameter. We use Markov chain Monte Carlo(MCMC) methods in our analysis. The Gibbs sampler is the most widely used MCMC methods which can be implemented in BUGS. There are several things that need to be considered for MCMC output: including the Monte Carlo error, the Gelman-Rubin convergence diagnostic and Deviance information criterion (DIC).

In the analysis of the MCMC output, *Monte Carlo error* (MC error) measures the variability of each estimate due to the simulation. MC error must be low in order to calculate the parameter of interest with increased precision. There are two common ways to estimate MC error: the *batch mean* method and the *windows estimator* methods. The first one is simple and easy to implement and has been used in WinBUGS. The batch mean estimator of the Monte Carlo error is discussed in details by Hastings (1970), Geyer(1992), Roberts (1996), Carlin and Louis (2000) and Givens and Hoeting (2005). Here we give a brief introduction to this method.

We partition the resulting output sample in K batches (K=30 or 50). The sample size of each batch $\nu = \tilde{T}/K$ must be sufficiently large in order to estimate the variance consistently and also eliminate autocorrelations. Before we calculate the Monte Carlo error of the posterior of $G(\boldsymbol{\psi})$,here $G(\boldsymbol{\psi})$ be arbitrary posterior, we first calculate each batch mean by

$$\overline{G(\boldsymbol{\psi})}_b = \frac{1}{\nu} \sum_{t=(b-1)\nu+1}^{b\nu} G(\boldsymbol{\psi}^{(t)}) \tag{3.8}$$

for each batch $b = 1, ..., K$, and the overall sample mean by

$$\overline{G(\boldsymbol{\psi})} = \frac{1}{\tilde{T}} \sum_{t=1}^{\tilde{T}} G(\boldsymbol{\psi}^{(t)}) = \frac{1}{K} \sum_{b=1}^{K} \overline{G(\boldsymbol{\psi})}_b \tag{3.9}$$

assuming that we keep $\boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(\tilde{T})}$ observations. The an estimate of the MC error is simply given by the standard deviation of the batch means estimates $\overline{G(\boldsymbol{\psi})}_b$

$$MCE[G(\boldsymbol{\psi})] = \sqrt{\frac{1}{K(K-1)} \sum_{b=1}^{K} \left(\overline{G(\boldsymbol{\psi})}_b - \overline{G(\boldsymbol{\psi})}\right)^2}$$

The Gelman-Rubin convergence diagnostic (Gelman and Rubin,1992) is available in WinBUGS via the *bgr diag* option. With multiple chains generating, each one starting from different initial values. Then an ANOVA-type diagnostic test is implemented by calculating and comparing the between-sample and the within-sample variability. The statistic R can be estimated by

$$\hat{R} = \frac{\hat{V}}{WSS} = \frac{\tilde{T}-1}{\tilde{T}} + \frac{BSS/\tilde{T}}{WSS}\frac{\kappa+1}{\kappa}$$

where κ is the number of generated sample/chains, $\tilde{T}$ is the number of iterations kept in each sample/chians, $BSS/\tilde{T}$ is the variance of the posterior mean values over all generated samples/chains(between sample variance), WSS is the mean of the variances within each sample(within-sample variability), and

$$\hat{V} = \frac{T^{'}-1}{T^{'}}WSS + \frac{BSS}{T^{'}}\frac{\kappa+1}{\kappa}$$

is the pooled posterior variance estimate. When convergence is achieved and the size of the generated data is large, then $\hat{R} \to 1$. Brook and Gelman (1998) adopted a corrected version of this statistic:

$$\hat{R}_c = \frac{d+3}{d+1}\hat{R}$$

where d is the estimated degrees of freedom for the pooled posterior variance estimate $\hat{V}$. We will use this statistic to check the convergence.

We use deviance information criterion (DIC) to compare different fitting models. We refer to Spiegelhalter *et al.* (2002) and Ntzoufras (2008, sec.6.4) for more details about DIC. The DIC is defined as $DIC = \bar{D} + p_D$ where $\bar{D}$ is the posterior mean of the deviance and p_D is the effective number of parameters in the model. The parameter p_D is calculated using $p_D = \bar{D} - D(\bar{\psi})$, where $D(\bar{\psi})$ is the deviance evaluated at the posterior mean of the unknown parameters. The DIC is particularly useful for complex hierarchical models where the numbers of parameters used is unknown. DIC is a generalization of the Akaike information criterion (AIC).

3.2 Model used in practice

Recall the formula 3.4, we start by ignoring frailties and setting $\lambda_0(t) = \lambda_0$, where λ_0 is baseline intensity parameter. Now we assume: the meaning of this assumption is that we have constant baseline hazard of default, all the time dependence comes through the macroeconomic covariates $\boldsymbol{X}_i(t)$. This model can be extended by adding a random effect b and allow both time-dependent and time-independent risk factors.

3.2.1 Model 3.1: Model with macroeconomic covariates

Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}(t)$ which are the same for all subjects and credit rating effect $\gamma_{r(i,t)}$ of the ith company at time t. The intensity function is given by:

$$\lambda_i(t|\boldsymbol{\gamma}, \boldsymbol{\eta}) = Y_i(t)\lambda_0 \exp(\boldsymbol{\eta}^{'}\boldsymbol{z}(t) + \boldsymbol{\gamma}_{r(i,t)}) \tag{3.10}$$

This model allows for difference in the events times that only depend on the macroeconomic variables and credit rating effect $\boldsymbol{\gamma}_{r(i,t)}$; it does not allow for the random effect.

3.2.2 Model 3.2: Model with yearly shared frailty

The model with macroeconomic covariates will be extended by adding a component of yearly heterogeneity $b_{y(t)}$ captures heterogeneity not captured by $\boldsymbol{z}(t)$, this frailty process is only depend on time t and will take the number yearly. Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}(t)$, credit rating effect $\boldsymbol{\gamma}_{r(i,t)}$ of the ith company at time t and yearly frailty. The intensity function is given by:

$$\lambda_i(t|\boldsymbol{\gamma}, \boldsymbol{\eta}) = Y_i(t)\lambda_0 \exp(\boldsymbol{\eta}^{'}\boldsymbol{z}(t) + \boldsymbol{\gamma}_{r(i,t)} + b_{y(t)}) \tag{3.11}$$

where $y(t)$ gives the year period corresponding to time t. This model allows for difference in the events times that depend on the macroeconomic variables and credit rating effect $\boldsymbol{\gamma}_{r(i,t)}$ as well as the shared frailty for each year. However, it does not allow serial dependence between each yearly shared frailty.

3.2.3 Model 3.3: Model with serial dependence for yearly shared frailty

The model with yearly shared frailty will be extended by adding serial dependence for yearly heterogeneity $b_{y(t)}$. Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}(t)$, credit rating effect $\boldsymbol{\gamma}_{r(i,t)}$ of the ith company at time t and yearly frailty. The intensity function is given by:

$$\lambda_i(t|\boldsymbol{\gamma}, \boldsymbol{\eta}) = Y_i(t) \exp(\boldsymbol{\eta}' \boldsymbol{z}(t) + \boldsymbol{\gamma}_{r(i,t)} + b_{y(t)}) \tag{3.12}$$

$$b_y = \alpha + \varphi b_{y-1} + \epsilon_y$$

A simple model AR(1) process has been used to capture the serial dependence for yearly shared frailty. This model allows for difference in the events times that depend on the macroeconomic variables and credit rating effect as well as yearly shared frailty with serial dependence.

3.2.4 Model 3.4: Model with shared sector frailties

We will add shared sector frailties as the second level of frailty. Then the frailties with two levels will become $b_{y(t),s(i)}$. The frailties will be $b_{y(t),s(i)}$ with yearly shared frailty extended by adding shared sector frailties. Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}(t)$, sector effect $s(i)$ of the ith company,rating effect $\boldsymbol{\gamma}_{r(i,t)}$ for ith company at time t and yearly frailty. The intensity function is given by:

$$\lambda_i(t|\boldsymbol{\gamma}, \boldsymbol{\eta}) = Y_i(t) \exp(\boldsymbol{\eta}' \boldsymbol{z}(t) + b_{y(t),s(i)} + \boldsymbol{\gamma}_{r(i,t)}) \tag{3.13}$$

$$b_{y,s} \sim N(b_y, \sigma^2)$$

$$b_y = \alpha + \varphi b_{y-1} + \epsilon_y$$

A simple model AR(1) process has been used to capture the serial dependence for yearly shared frailty. This model allows for difference in the events times that depend on the macroeconomic variables and sector effect as well as two levels of shared frailties with serial dependence. We can tell the difference for companies in different sectors in different year.

3.3 Empirical study of default data

3.3.1 Data description

We use the Standard & Poor's database CreditPro 6.6 which consists of 10439 companies from 13 industry sectors over the period January 1981 to December 2003. Overall 19054 effective rating migrations are recorded in the CreditPro database as well as 1386 defaults. Here we remove the rating transitions which started with rating category "NR" and treat the transitions as censored for those ending with "NR".

The rating classes included in this analysis are

$$\mathcal{K} = \{CCC, B, BB, BBB, A, AA, AAA, D\}$$

where we merge actual rating k^+, k, k^- into k. We also merge CCC, CC, and C into a single rating class CCC. The whole database consists of 13 industries among 93 countries or region and 6897 of them are US companies. The industry sectors in this analysis are modified from the S&P industry sectors. We merged 13 industries into 8 industries, Consumer/Service with Transportation, Energy with Utility, Financial Institution with Insurance and Real Estate, Forest and Building products with Homebuilders, Health Care with Chemicals and finally High Technology with Telecommunications. The industry sectors we merged are have similar business which have high correlation between industry sectors.

We start to look at default analysis of the whole database in this analysis in order to study the models suggested. Rating class AAA and AA which rarely default have

Sector	Name	#Obligors
1	“Aerospace/automotive/capital goods/metal”	803
2	“Consumer/service sector”+“Transportation”	1238
3	“Energy and natural resources”+“Utility”	903
4	“Financial Institutions”+“Insurance”+“Real estate”	1340
5	“Forest and building products”+“Homebuilders”	258
6	“Health care”+ “Chemicals”	448
7	“Hightec/computers/office equipment”+“Telecom”	588
8	“Leisure time/media”	644

Table 3.1: Numbers of obligors used for different industry sectors

been excluded from this study, they will be reconsidered in our transition analysis in the following chapters. Here we consider the rating $\mathcal{K}_d = \{CCC, B, BB, BBB, A\}$ to default D of US companies which include 6222 companies as well as 1356 defaults (some companies censored). The details about the industries for the data we used in this analysis are shown in Table 3.1.

3.3.2 Results

There are several ways to implement Gibbs simulation. WinBUGS is one of the most popular ways to implement Gibbs simulation. In running the Gibbs sampling algorithm, the prior specifications are very important, we set them up as follows: for all the constant baseline, the precision τ of time-dependent fixed effect macroeconomic variables was set to 0.001, resulting a normal distribution which is very uninformative. A non-informative Gamma prior is assumed for τ, the precision of the frailty parameters. Note that the above ‘additive’ formulation of the frailty model is equivalent to assuming multiplicative frailties with a log-Normal population distribution. Clayton (1991) discusses the Cox proportional hazards model with multiplicative frailties, but assumes a Gamma population distribution.

Model 3.1: Model with macroeconomic covariates

The rating migration or default depend on the "state-of-the-economy". There are several possible proxies and we have shown that Chicago Federal National Activities Index three month moving average (CFNAIMA3) is the best macroeconomic covariate to capture the US economy cycle among them, therefore we use the CFNAIMA3 as the macroeconomic covariates in this analysis. All the details about macroeconomic covariates can be found in Chapter 2. Obviously we could easily include further macroeconomic covariates if desired, but CFNAIMA3 captures the main business cycle effect.

Model 3.1: Model with macroeconomic covariates only							
	μ	η	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}
$mean$	-3.625	-1.225	-7.333	-5.428	-4.109	-2.656	0
sd	(0.0416)	(0.0437)	(0.5288)	(0.2372)	(0.1301)	(0.0658)	0
$MCerror$	0.0010	0.0008	0.0070	0.0039	0.0020	0.0012	0

Table 3.2: Model 3.1 with macroeconomic variable only

With 10000 iterations and burn-in 5000, the first 5000 iterations were discarded. The posterior estimation for Model 3.1 are given in Table 3.2. The posterior mean of macroeconomic covariate CFNAIMA3 coefficient η is -1.225 and standard deviation(SD) 0.04372, which means the increase of CFNAIMA3 will reduce the default intensity for all the rating category. The posterior mean of μ is -3.625 with standard deviation 0.04162, which is the parameter for constant baseline, $\exp(\mu)$ is the constant baseline in our model, and $\exp(\mu+\boldsymbol{\eta}'\boldsymbol{z}(t))$ equal to baseline including macroeconomic effect. The rating category is treated as a fixed effect in our analysis, we set the default from rating CCC to 0, and the rest coefficients of rating category shows that higher rating with lower default intensity. The MC error number is quite small which means the model fits the data well. Figure 3.1 compares log-scale intensity function for all ratings.

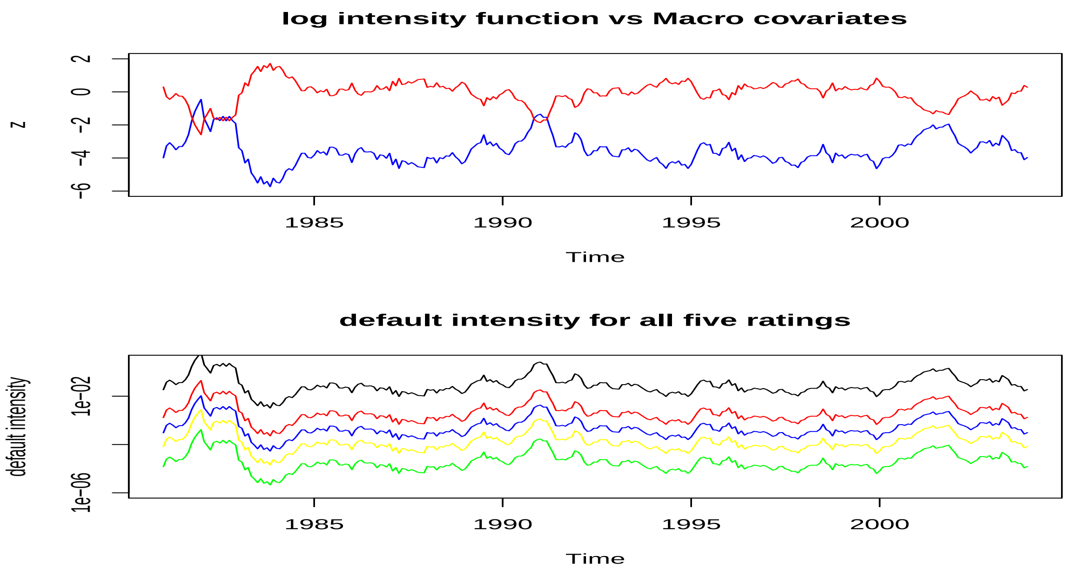

Figure 3.1: Intensity vs macroeconomic covariates in upper figure and monthly intensity for all five different rating categories with Model 3.1

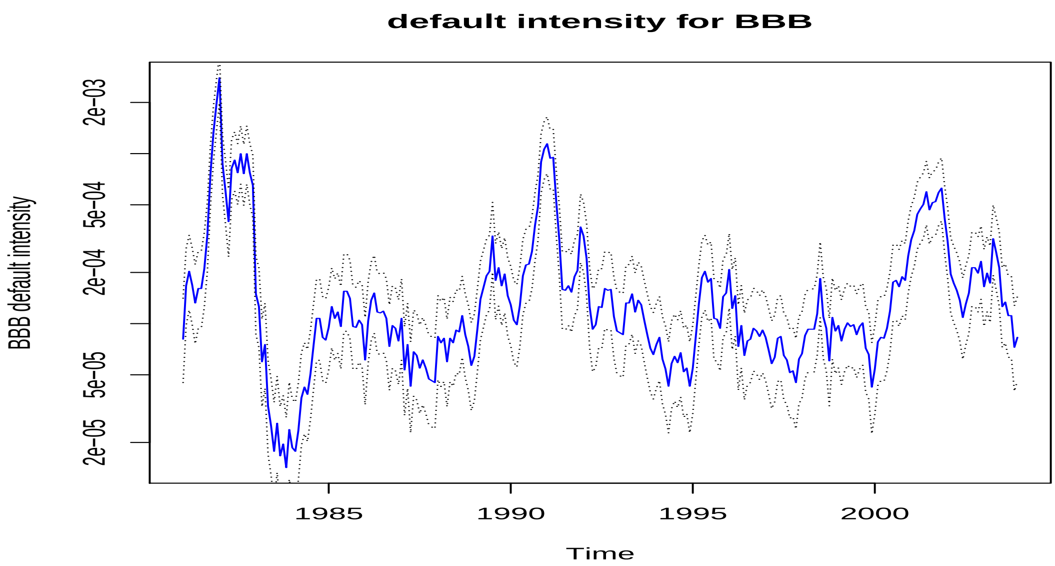

Figure 3.2: Monthly intensity for rating category BBB with Model 3.1; posterior means with 95% credible intervals

Model 3.2: Model with yearly shared frailty								
	μ	η	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}	σ
mean	-2.699	-0.3981	-7.307	-5.411	-4.085	-2.617	0	0.919
sd	0.2098	0.0897	0.5296	0.2375	0.1316	0.0661	0	0.1548
MCerror	0.0203	0.0032	0.0087	0.0037	0.0022	0.0014	0	0.0038

Table 3.3: Model 3.2 with yearly shared frailty

The upper figure shows opposite direction of the monthly intensity and macroeconomic covariates CFNAIMA3 which confirm the negative effect of macroeconomic effect for intensity of default rating. The increase of CFNAIMA3 will reduce the default intensity for all the credit rating categories in this analysis. The lower figure show that lower rating has higher default intensity. The lowest log-intensity to the highest are represent rating categories A, BBB, BB, B and CCC to default respectively. Figure 3.2 shows the estimated monthly default intensity for rating category BBB and its corresponding 95% credible intervals. For simplicity, we will show other rating category's default intensity in the following models. The monthly intensity for rating category BBB changes over time and shows opposite direction of macroeconomic covariate.

Model 3.2: Model with yearly shared frailty

The rating migration or default depend on the "state-of-the-economy". However, further random effect is needed to capture patterns of variability in responses that cannot be explained by the observed macroeconomic covariates alone. Random effects are additional unobserved factors help to explain the rating migration activity. In Model 3.2, we add yearly shared frailties and try to capture the variability between different years.

As in Model 3.1, we run 10000 iterations and burn-in 5000, the first 5000 iterations were discarded. The results for Model 3.2 are presented in Table 3.3. The MC error number for all parameters are quite small which means the model fits the data well and results are acceptable. The posterior mean for baseline hazard effect μ is increased from -3.625 in Model 3.1 to -2.699 in Model 3.2, with σ the standard deviation 0.2098, which is the parameter for constant baseline, $\exp(\mu)$ is the constant baseline in this model, and $\exp(\mu+\boldsymbol{\eta}'\boldsymbol{z}(t))$ equal to baseline including macroeconomic effect. The rating fixed effect shows that lower rating has higher intensity function. The systematic risk factors in Model 3.1 can be divided into observed fixed effects and unobserved yearly shared frailty to capture heterogeneity for default intensity. We compared the observed factor Model 3.1 with the unobserved factor in Model 3.2 and will show it in the default intensity. There is some co-movement between these two series. The observed macroeconomic covariate does capture the credit rating transitions but not fully capture them. In Figure 3.3, we display default intensity for rating category B, two risk factors $\exp(\mu+\boldsymbol{\eta}'\boldsymbol{z}(t))$ in Model 3.1 and $\exp(\mu+\boldsymbol{\eta}'\boldsymbol{z}(t)+b)$ in Model 3.2 are shown. The risk factors in Model 3.2 generally have higher default intensities than in Model 3.1. The unobserved yearly shared frailty in Model 3.2 to capture heterogeneity which cannot be explained by observed macroeconomic covariates. The macroeconomic covariates around 1990s affect default intensity in Model 3.1 much more than in Model 3.2 since it is the only factor in Modle 3.1. Figure 3.4 shows the estimated monthly intensity for rating category B and its corresponding 95% credible intervals.

Model 3.3: Model with serial dependence for yearly shared frailty

In Model 3.2, the realizations of yearly shared frailty b_y are assumed to be independent over different time periods (yearly in our analysis). We can extend this model by adding a dependence structure on the yearly shared frailty. In credit risk modelling, it seems appropriate to assume the current value of yearly shared frailty b_y depends on its previous time period. We introduce a first-order autoregressive, AR(1), time series. Here, we remove the baseline coefficient μ and introduce a mean α for autoregressive process AR(1). The baseline hazard can be calculated by $\mu = \frac{\alpha}{1-\phi}$.

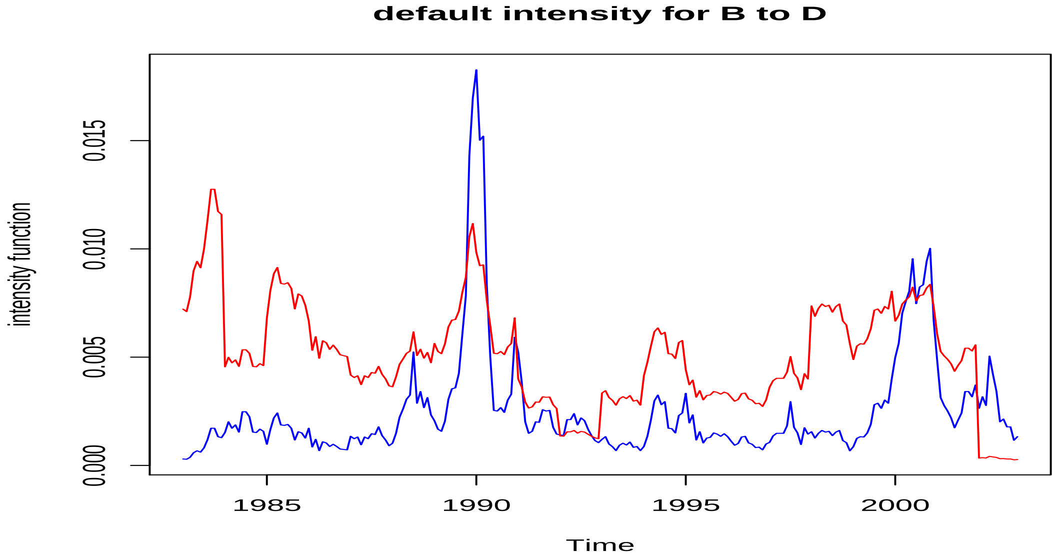

Figure 3.3: Default intensity for rating category B: The blue plot displays the intensity for Model 3.1 with macroeconomic only and the red plot display the intensity with unobserved yearly shared frailty b in Model 3.2

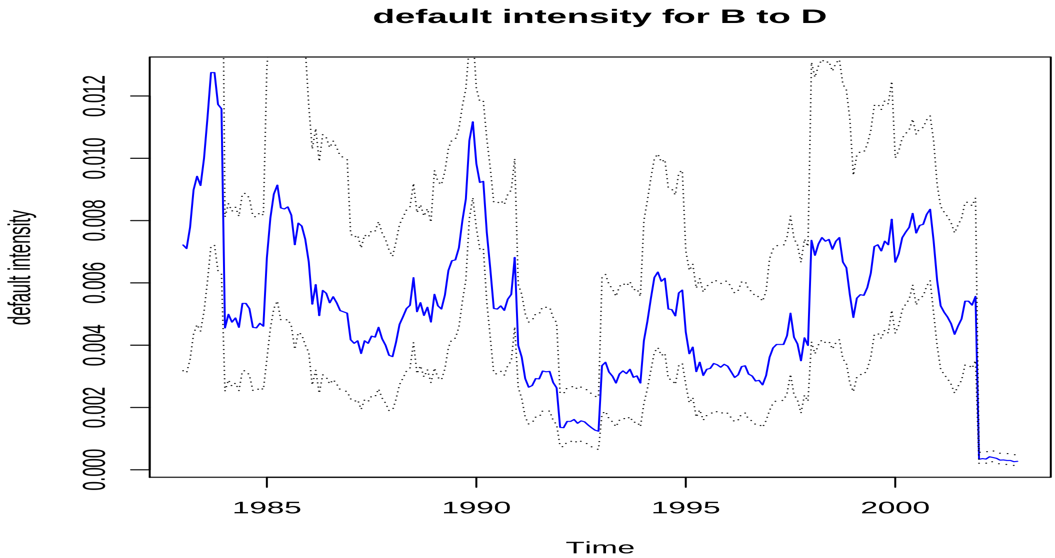

Figure 3.4: Monthly default intensity for rating categories B with Model 3.2; posterior means with 95% credible intervals

Model 3.3: Model with serial dependence for yearly shared frailty							
node	α	η	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}
mean	-1.167	-0.3889	-7.323	-5.420	-4.091	-2.616	0
sd	0.6962	0.0911	0.5320	0.2356	0.1324	0.0660	0
MCerror	0.0225	0.0039	0.0078	0.0037	0.0021	0.0013	0

Remaining Parameters		
	σ	ϕ
mean	0.7710	0.6139
sd	0.1383	0.2692
MCerror	0.0028	0.0089

Table 3.4: Model 3.3 with serial dependence for yearly shared frailty

We run 10000 iterations and burn-in 5000, the first 5000 iterations were discarded. The results for Model 3.3 are presented in Table 3.4. The MC error number for all parameters are quite small which means the model fits the data well and results are acceptable. The posterior mean for autoregressive coefficient α is -1.167 with standard deviation 0.6962 and ϕ is 0.6139 with standard deviation 0.2692. So the baseline hazard coefficient μ is -3.0225 where μ is -2.699 in Model 3.2. The macroeconomic covariate effect η increase from -0.3981 to -0.3889, so the total number is almost the same in Model 3.2 and Model 3.3. The variance of frailty has posterior mean 0.8725 and standard deviation 0.1636, which means apart from the serial dependence, the event time within a time period share a common frailty effect that partially summarizes the dependence within the time period.

In Figure 3.5, we display default intensity for rating category BB, three risk factors $\exp(\mu+\boldsymbol{\eta}'\boldsymbol{z}(t))$ in Model 3.1 and $\exp(\mu+\boldsymbol{\eta}'\boldsymbol{z}(t)+b)$ in Model 3.2 and Model 3.3 shows co-movement. The difference between Model 3.2 and Model 3.3 is the latter include a serial dependence AR(1) process. The unobserved yearly shared frailty in Model 3.2 and Model 3.3 to capture heterogeneity which cannot be explained by observed macroeconomic covariates. All the three lines share the co-movement, where the red

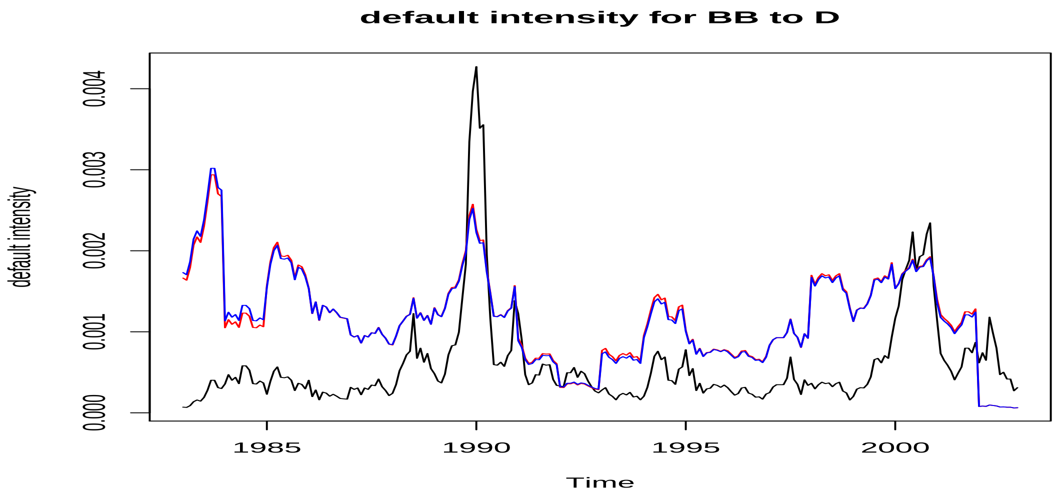

Figure 3.5: Default intensity for rating category BB. The black plot displays the default intensity for Model 3.1 with macroeconomic only and default intensity function with unobserved yearly frailty b in Model 3.2 (red line) and Model 3.3 (blue line)

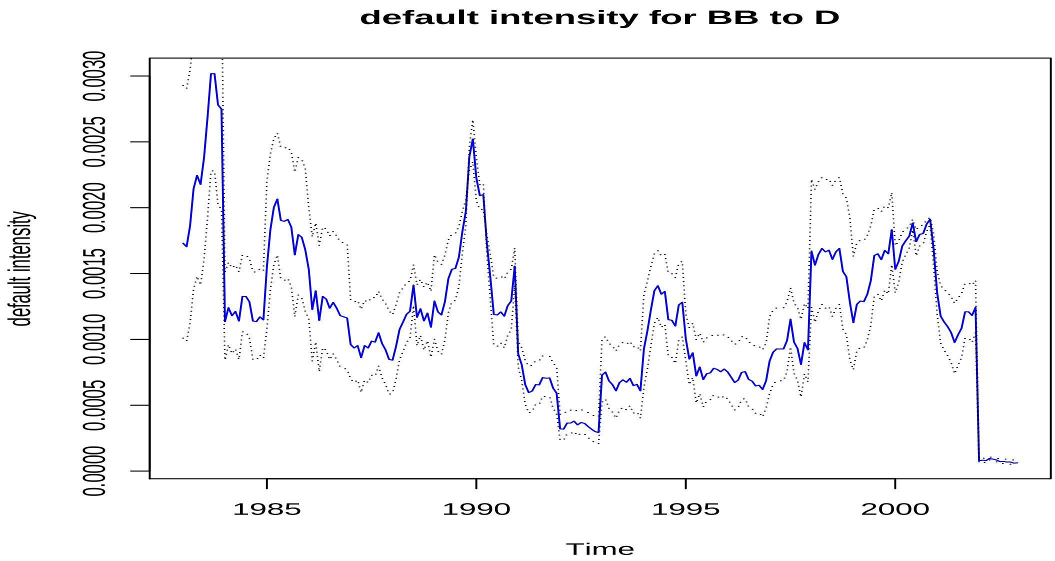

Figure 3.6: Monthly default intensity for rating categories BB with Model 3.3; posterior means with 95% credible intervals

Model 3.4: Model with shared sector frailties

node	α	η	γ_A	γ_{BBB}	γ_{BB}	γ_B	γ_{CCC}
mean	-1.135	-0.3847	-7.404	-5.492	-4.183	-2.728	0
sd	0.6581	0.09262	0.537	0.2362	0.1353	0.06929	0
MCerror	0.02063	0.00501	0.00814	0.00402	0.00288	0.00218	0

Remaining Parameters

	ϕ	σ	σ_1
mean	0.6264	0.747	0.4303
sd	0.2563	0.1467	0.06121
MC.error	0.00539	0.00203	0.00193

Table 3.5: Model 3.4: Model with shared sector frailties (have been omitted for a simpler presentation)

line (Model 3.2) and blue line (Model 3.3) are almost identical. The serial dependence shows within a time period default share a common frailty. The adding of AR(1) process improves not as significant as the introducing of yearly shared frailties. The risk factors in Model 3.3 generally have higher default intensities than in Model 3.1 but lower default intensities in particular time period which have peak macroeconomic covariates. Figure 3.6 shows the estimated monthly intensity for rating category BB and its corresponding 95% credible intervals.

Model 3.4: Model with shared sector frailties

We use yearly yearly shared frailty to capture the variability for different time period for any company. Further frailties like industry sector can be introduced to capture additional variability. We divided the CreditPro data into 8 industry sectors.

We run 10000 iterations and burn-in 5000, the first 5000 iterations were discarded. The part of results for Model 3.4 are presented in Table 3.5 which shows two levels of frailties. The MC error number for all parameters are quite small which means

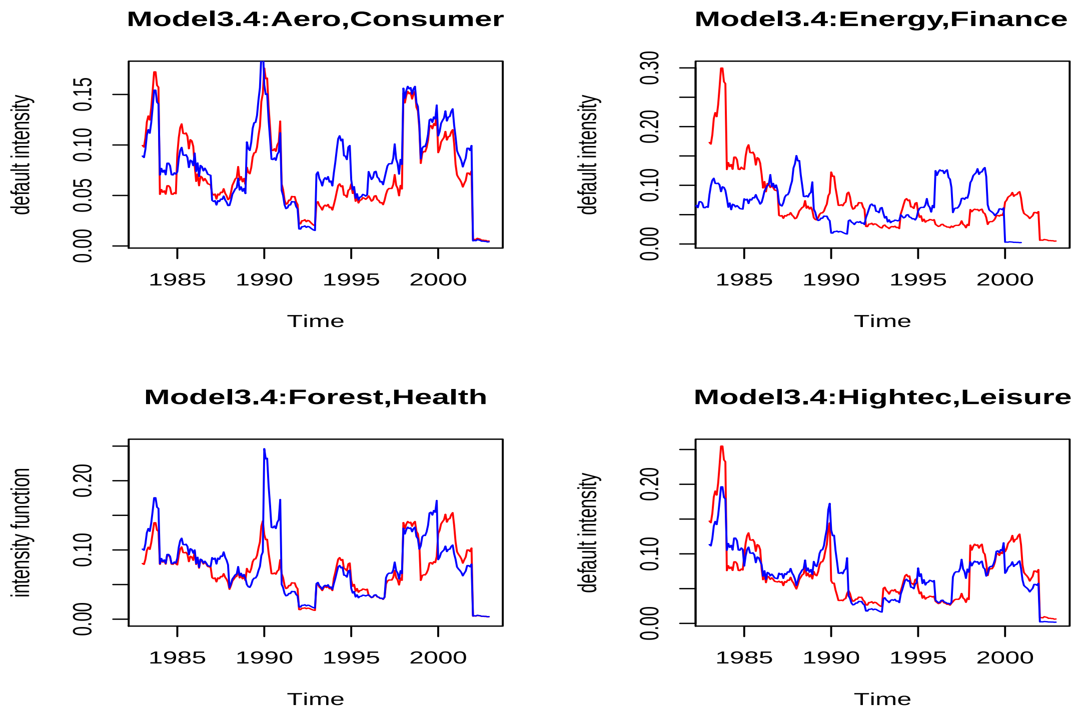

Figure 3.7: Different sector's default intensity for rating category CCC

the model fits the data well and results are acceptable. The posterior mean for autoregressive coefficient α is -1.135 with standard deviation 0.6581 and ϕ is 0.6264 with standard deviation 0.2563. So the baseline hazard coefficient μ is -3.0380 where μ is -3.0225 in model 3.3. The macroeconomic covariate effect η decrease from -0.3889 to -0.3847, but the intensity keep same in model 3.3 and model 3.4.

The following Figure 3.7 shows default intensity for rating category CCC with 8 different industry sectors in Model 3.4 and try to show the co-movements and heterogeneity for 8 different industries.

We can easily find the co-movement for each sector as well as the heterogeneity. The Aero, Consumer, Forest and Health show higher default intensity around 1990 but Energy and Finance shows lower default intensity at the same time. Almost all the industry sectors shows lower default intensity around 1995. With introducing the second-level shared sector frailties, it capture the heterogeneity which cannot be ex-

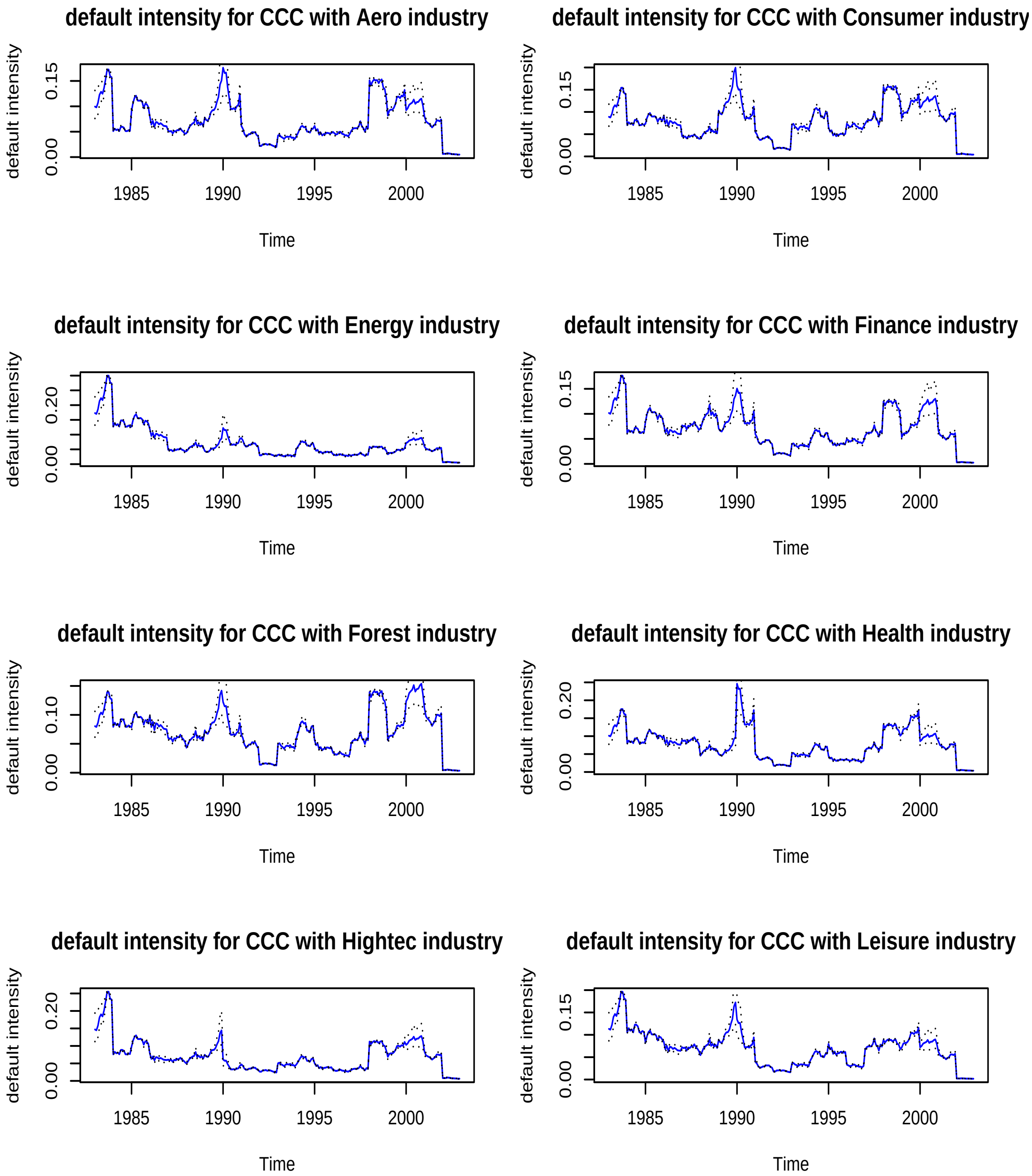

Figure 3.8: Monthly default intensity for rating categories CCC with Model 3.4; posterior means with 95% credible intervals

	Dbar	Dhat	pD	DIC
Model 3.1	12089.300	12083.400	5.933	12095.200
Model 3.2	10956.500	10930.200	26.339	10982.900
Model 3.3	10954.500	10928.600	25.956	10980.500
Model 3.4	10822.600	10736.900	85.674	10908.300

Table 3.6: Deviance information criterion for four different Models

plained by the one-level yearly shared frailty model. Figure 3.8 shows the estimated monthly intensity for rating category CCC and its corresponding 95% credible intervals for 8 different industry sectors.

Deviance information criterion (DIC)

We use deviance information criterion (DIC) to compare these different models.

For each model discussed above, the Gibbs sample was run for 10000 iterations. The first 5000 iterations were discarded, and the remaining 5000 iterations were used for analysis. The estimates for DIC and p_D for the four different models are presented in Table 3.6. The use of DIC is rather to compare different models than choose the true model. Model 3.2 and 3.3 have introduced yearly shared frailty and these are significantly better than Model 3.1. With AR(1) for yearly frailty, Model 3.3 is sightly better than Model 3.2 without AR(1) from the DIC estimation. By introducing second level of shared sector frailties,, Model 3.4 improved compared with Model 3.2 and Model 3.3. The more complexity of model we introduced in this analysis, the better performance according the DIC.

3.4 Discussion

In this chapter, we apply Manda and Meyer(2005) model for default risk model and allow two-levels shared frailties which are yearly shared frailty and shared sector frailties. Survival model with shared frailties allow the systematic portfolio risk to be divided into observed fixed effects and unobserved random effects to capture heterogeneity in rating default risk. The shared sector frailties model can even capture heterogeneity in time and across industry sector or other different groups like country. Because CFNAIMA3 capture the main business cycle effect, therefore we did not include other macroeconomic covariates in this analysis. For implement the survival frailty model, MCMC is one of the best technique could be used for model calibration. We chose WinBUGS to implement our model. Using deviance information criterion (DIC), we show that a model with serial dependence and two-levels of shared frailties, which accounts the heterogeneity in both time period and industry sectors , provides a better fit than other models.

However, only the default model was considered in this study, in order to make our model universally applicable for credit modelling. We will extend our model from default risk to rating transition. We exclude the highest rating AAA and AA which rarely default and non-US companies from our modelling. We will reconsider these two ratings in our transition risk model.

Chapter 4

Modelling migration risk with survival models

From the definition of credit risk, we know it is the risk of the change in value of a portfolio caused by unexpected changes in the credit quality of issues or trading partners. Default risk is a special case of transition risk. In Chapter 3, we used time-to-event methods in the survival framework to model the default risk and would like to extend the models in chapter 3 to general models for all transition risk.

We extend a simpler model used in medical statistics (Manda and Mayer (2005)) for the credit risk application. We estimate rating transition model with shared dynamic frailties for different industry sectors and macroeconomic covariates using Bayesian techniques (MCMC). In this chapter each transition intensity follows a Cox type multiplicative regression model with two levels of frailties to account for both time period and industry sector heterogeneity. Delloye, Fermanian and Sbai (2006) define a reduced-form credit portfolio model which treat rating transition as independent competing risks with conditionally independent and proportional hazards assumption. They also allow strong dependence levels by adding heterogeneity. We consider the whole transition data and let all rating transitions share the same macroeconomic covariates and unobservable random process for all companies in monthly interval although Delloye, Fermanian and Sbai (2006) split the Standard & Poors' data into several groups based on the similar rating transition type and analyze these separately.

We have five aims in this chapter and the following chapter: First, we will extend analysis of Chapter 3 to all transitions with Standard & Poor's CreditPro data. Second, we will consider a simplification in which we model the size of transition measured by the number of notches on a rating scale. Third, we will derive genuine point-in-time (PIT) transition matrices. Fourth, we will look at the sectoral variation. Fifth, we will show how a Bayesian Gibbs sampling solution is possible. We will extend Cox's hazard model which has been increasingly used to model the hazard of credit risk events in recent years. We will consider randomness in the intensity function caused by shared frailties that are associated with unobserved business cycle covariates. Such a model can be analysed in Anderson & Gills counting process formulation. In this chapter, firms are subject competing hazards because one ratings could go to different rating at the end of time period.

In this chapter, we will briefly introduce the Manda and Meyer (2005) model and its estimation methods. Then we will apply Manda and Meyer (2005) with time-to-event credit risk models with credit transition data. We will present application to Standard & Poor's CreditPro data with time-dependent frailty model for recurrent failure time data in Bayesian context and estimate it using the Markov Chain Monte Carlo methods. McNeil and Wendin (2006) use Bayesian techniques for ratings migration modelling. As we discussed in the previous chapter, Bayesian methods based on Markov Chain Monte Carlo (MCMC) techniques have three advantages: Bayesian methods improves the estimation accuracy, Bayesian estimation also allows for taking into account expert opinion through the use of subjective prior distribution of for model parameters, and Bayesian inference becomes straightforward to compute the default and transition probabilities. We will provide comparative results with credit risk models using GLMMs. The random effect in GLMMs models will appear as frailties in survival models. An autoregressive process is used as time-dependent frailty and two levels frailty time and industry will be used in this model as in default model. We will investigate differences in transition intensity between industry sectors. In this chapter, we will model the frailty models for credit transitions by numbers of levels (notches) and actual credit rating transitions.

4.1 Theory

Anderson and Gill (1982) extend the Cox model to a model where covariate processes have a proportional effect on the intensity process of a multivariate counting process. We extend the Anderson-Gill model in chapter 3 to allow multiple event types which are suitable for modelling multiple transitions.

4.1.1 Models

Our model is relative to that of Manda and Meyer (2005) in medical research. We apply these ideas to credit rating transitions for the first time. As in default model, two-levels of frailties for credit risk modelling is allowed. At every time t, any company i has k possible ratings categories. The time durations are assumed independent conditionally on the macro-economic process and the idiosyncratic firm characteristics. Time of events are ordered $0 < T_{i1} < T_{i2} < \cdots$, suppose the study period for company i is partitioned into finite disjoint intervals $A_t = [t, t+dt), 0 \leq t \leq c_i$ such that no time interval contains more than one of the consecutive events (rating transitions) of the company. The predetermined time interval $[0, c_i]$ is divided into discrete time intervals. In practice monthly intervals prove to be sufficiently small. Therefore we finally choose monthly for our analysis. Thus, the company's counting process jump for transition couple (h, j)

$$dN_{hji}(t) = N_{hji}(t+dt) - N(t)$$

takes only value 0 and 1. $N_{hji}(t)$ is an observable process that counts the number of events occurred up to time t for transition (h, j); $Y_{hi}(t)$ is a non-negative predictable process taking the value 1 if the company i is under observation and the value 0 otherwise. $\boldsymbol{X}_i(t)$ is a d-dimensional vector of risk factors, including time-dependent macro-economic process and the time-independent fixed effects. For i^{th} subject, the corresponding intensity at time t is $\lambda_{hji}(t|\boldsymbol{\beta})$ for subject i moving from start state h to j. As in chapter 3, the frailty-based Andersen-Gill frailty model will be extended by adding shared frailty $W_{hji}(t)$ which can capture the risk not accounted by observed risk variables for any transition from h to j. The intensity function (3.2) and (3.3) will be extended as following for multiple events: This implies that

$$\lambda_{hji}(t|\boldsymbol{\beta})dt = P(dN_{hji}(t) = 1|F_{t-}; \boldsymbol{\beta}, W_{hji}(t))$$

where F_{t-} is the available data just before time t. The intensity function for a transition type (h, j) at time t for company i is given by

$$\lambda_{hji}(t|\boldsymbol{\beta}) = Y_{hi}(t)\lambda_{hj0}(t)\exp(\boldsymbol{\beta}'_{hj}\boldsymbol{X}_i(t) + W_{hji}(t)) \tag{4.1}$$

where $\lambda_{hj0}(t)$ is baseline intensity function. And $\boldsymbol{\beta}_{hj}$ is a parameter for each transitions (h, j). As in the default model, this model can be extended by adding random effect $W_s(i)$ and allow both time-dependent and time-independent risk factors.

The Andersen-Gill model with dynamic frailty in (3.4) will be extended as following for multiple events:

$$\lambda_{hji}(t|\boldsymbol{\beta}, W_{hji}(t)) = Y_{hi}(t)\lambda_{hj0}(t)\exp(\boldsymbol{\beta}'_{hj}\boldsymbol{X}_i(t) + W_{hji}(t)) \tag{4.2}$$

4.1.2 Parameter estimation

We have shown the parameter estimation methods for Cox and Andersen-Gill model in chapter 3. Here we will extend the full maximum likelihood methods and Bayesian inference in order to fit the transition data. Bayesian methods based on Markov Chain Monte Carlo (MCMC) techniques have three advantages: Bayesian methods improves the estimation accuracy, Bayesian estimation also allows for taking into account expert opinion through the use of subjective prior distribution of for model parameters, and Bayesian inference becomes straightforward to compute the default and transition probabilities. WinBUGS can use either a standard 'point-and-click' windows interface for controlling the analysis, or can construct the model using a graphical interface called DoodleBUGS. R2WinBUGS is a package running WinBUGS from R. Using this package, it is possible to call a BUGS model, summarize inferences and convergence in a table and graph, and save the simulations in arrays for easy access in R.

Let $\lambda_{hji}(t)$ be the intensity for firm i start with state h and end with state j. In formula (4.2), for each transition $(h, j), h \neq j$, $\boldsymbol{\beta}$ is the parameter to be estimated and $\boldsymbol{X}$ is the macroeconomic covariates in this analysis. In our simple model, covariate does not depend on each transition (h, j) but could extend to transition dependent by considering different explanatory macroeconomic variables for different transition transition (h, j). The observable process $N_{hji}(t)$ which denote the number of transitions for firm i from h to j at period time t. $Y_{hi}(t)$ be the indicator variable which take value 1 or 0 depend on under observation or not.

The likelihood in (3.7) is extended by the following:

$$\mathcal{L} = \prod_{i=1}^{n} \mathcal{L}_i$$

$$\mathcal{L}_i = \left\{ \prod_t \prod_{j \neq h} (\lambda_{hji}(t|X))^{dN_{hji}(t)} \right\} \cdot \left(- \sum_{j \neq h} \int_0^{\infty} Y_{hi}(u) \lambda_{hji}(u|X) du \right) \tag{4.3}$$

The log-likelihood will be:

$$\ln \mathcal{L}_i = \sum_{i=1}^{n} \ln \mathcal{L}_i$$

$$\begin{aligned}
\ln \mathcal{L}_i &= \sum_t \sum_{j \neq h} dN_{hji}(t)[\ln \lambda_{hj0} + \boldsymbol{\beta}'_{hj} \boldsymbol{X_i}(t) + W_{hji}(t)] - \sum_{j \neq h} \int_0^{\infty} Y_{hi}(u) \lambda_{hj0}(u) \exp(\boldsymbol{\beta}'_{hj} \boldsymbol{X}(u) + W_{hji}(u)) du \\
&= \sum_{j \neq h} \left\{ \sum_t dN_{hji}[\ln \lambda_{hj0}(t) + \boldsymbol{\beta}'_{hj} \boldsymbol{X_i}(t) + W_{hji}(t)] - \int_0^{\infty} Y_{hi}(u) \lambda_{hj0}(u) \exp(\boldsymbol{\beta}'_{hj} \boldsymbol{X}(u) + W_{hji}(u)) du \right\} \\
&= \sum_{j \neq h} \ln \mathcal{L}^*_{hji}
\end{aligned}$$

$$\ln \mathcal{L} = \sum_{j \neq h} \ln \mathcal{L}^*_{hj}$$

This likelihood can be split into a sum of each transition (h, j) which means we can estimate the parameters separately. This is very useful and will reduce the implementations in practice.

We extend Manda and Meyer (2005) model to use transition data. As in chapter 3 , the Bayesian inference using time-dependent frailty will be discussed. The joint distribution of the observed data D given $\lambda_{hj0}, \boldsymbol{\beta}_{hj}, W$ is given by

$$\begin{aligned}
p(D|\lambda_{hj0}, \boldsymbol{\beta}_{hj}, W_{hji}(t)) &= \prod_{i}^{I} \left\{ \prod_{t\geq 0}^{c_i} \prod_{j\neq h} [Y_{hi}(t) e^{\boldsymbol{\beta}'_{hj} \boldsymbol{X}_i(t) + W_{hji}(t)} \lambda_{hj0}(t)]^{dN_{i(t)}} \right\} \\
&\times \exp\left(-\sum_{j\neq h} \int_0^{c_i} Y_{hi}(t) e^{\boldsymbol{\beta}'_{hj} \boldsymbol{X}_i(t) + W_{hji}(t)} \lambda_{hj0}(t) dt \right) \\
&\propto \prod_{i}^{I} \prod_{j\neq h} \left\{ \prod_{t\geq 0}^{c_i} [e^{\boldsymbol{\beta}'_{hj} \boldsymbol{X}_i(t) + W_{hji}(t)} d\Lambda_{hj0}(t)]^{dN_{i(t)}} \right. \\
&\times \left. \exp(-Y_{hi}(t) e^{\boldsymbol{\beta}'_{hj} \boldsymbol{X}_i(t) + W_{hji}(t)} d\Lambda_{hj0}(t)) \right\}
\end{aligned}$$

There are several ways to compare the performance of different models with posterior MCMC output: including the Monte Carlo error and The Gelman-Rubin convergence diagnostic and Deviance information criterion (DIC).

4.2 Frailty models for credit transitions by numbers of levels (notches)

Transition matrices are at the center of modern credit risk management. The reports on rating migrations published by Standard and Poors and Moodys are studied by credit risk managers everywhere and several of the most prominent risk management tools are built around estimates of rating migration probabilities. Transition matrices are widely used for risk management purpose, economic capital purpose and credit derivatives pricing. In this section, we will extend the default model in previous chapter and build the transitions model with different numbers of levels (notches) with frailty model. Then an empirical study using Standard & Poor's CreditPro data will be provided and discussed.

4.2.1 Theoretical models for credit rating transitions by numbers of levels (notches)

In this section we extend credit default models to credit rating transitions by numbers of levels (notches). For rating migration to its neighbour rating called 1-notch upgrade or downgrade, however 1-notch upgrade and downgrade are two different kinds of type rating transition. This model does make strong assumption which is that the rating transition of n-notches is the same type of event regardless whether the start rating is lower or higher. The multi-state feature of the model is represented as a set of transition types, $\mathbb{U} = \{1, 2, \ldots, U\}$. We consider all the possible n-notches in our sample database.

Model 4A.1: Model with macroeconomic covariates only

Suppose the intensity function depends only on time-dependent macroeconomic covariates $\boldsymbol{z}_t$ and transition type $u \in \mathbb{U}$ of the ith company. The monthly intensity function is given by:

$$\lambda_{ui}(t|\boldsymbol{\eta}, \upsilon) = Y_i(t) \exp(\boldsymbol{\eta}_u^{'} \boldsymbol{z}(t) + \upsilon_u) \tag{4.4}$$

The intensity function only dependent on macroeconomic covariates and transition type u. Each transition type u has different effect with macroeconomic covariates. This model allows for difference in the events times that only depend on the macroeconomic variables for each transition type; frailty is not considered in this simple model. With our strong assumption, rating transition of n-notches is the same kind of event.

Model 4A.2: Model with yearly time-dependent shared frailty

The model with macroeconomic covariates is extended by adding a component of yearly heterogeneity $b_{y(t)}$, this frailty process is only depend on time t for each transition type u and will take the number yearly. And we suppose each transition of n-notches has its own yearly heterogeneity. Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}_t$, transition of n-notches effect υ_u, γ_u and yearly frailty. The monthly intensity function is given by:

$$\lambda_{ui}(t|\boldsymbol{\eta}, \upsilon, \boldsymbol{\gamma}) = Y_i(t) \exp(\boldsymbol{\eta}_u^{'} \boldsymbol{z}(t) + \upsilon_u + \boldsymbol{\gamma}_u b_{y(t)}) \tag{4.5}$$

where $y(t)$ gives the year period where t belongs to. For each rating transition type u, macroeconomic covariates have different parameter η and different frailty process b_y. This model allows for difference in the events times that depend on the macroeconomic variables and transition of n-notches as well as the random effect for each year for each transition type. However, it does not allow serial dependence between each yearly random effect.

Model 4A.3: Model with serial dependence for yearly shared frailty

The model with yearly shared frailty will be extended by adding serial dependence for yearly heterogeneity $b_{y(t)}$. Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}_t$, transition of n-notches υ_u, $\boldsymbol{\gamma}_u$ and yearly shared frailty. The monthly intensity function is given by:

$$\lambda_{ui}(t|\boldsymbol{\eta}, \upsilon, \boldsymbol{\gamma}) = Y_i(t) \exp(\boldsymbol{\eta}_u^{'} \boldsymbol{z}(t) + \upsilon_u + \boldsymbol{\gamma}_u b_{y(t)}) \tag{4.6}$$

$$b_y = \varphi b_{y-1} + \epsilon_y$$

A simple model AR(1) process has been used to capture the serial dependence for yearly shared frailty. This model allows for difference in the events times that depend on the macroeconomic variables and transition of n-notches as well as random effect with serial dependence.

Model 4A.4: Model with shared sector frailties

We will add sector frailty as the second level of frailty. Then the frailty with two levels will become $b_{y(t),s(i)}$, $s(i)$ denotes sector of obligor i. The model with yearly shared frailty will be extended by adding serial dependence for yearly heterogeneity $b_{y(t)}$. Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}_t$, transition of n-notches and yearly, sector two levels of frailty. The monthly intensity function is given by:

$$\lambda_{u,i}(t|\boldsymbol{\eta}, \upsilon, \boldsymbol{\gamma}) = Y_i(t)\exp(\boldsymbol{\eta}_u^{'}\boldsymbol{z}(t) + \upsilon_u + \boldsymbol{\gamma}_u b_{y(t),s(i)}) \tag{4.7}$$

$$b_{y,s} \sim N(b_y, \sigma^2)$$

$$b_y = \varphi b_{y-1} + \epsilon_y$$

A simple model AR(1) process has been used to capture the serial dependence for yearly shared frailty. This model allows for difference in the events times that depend on the macroeconomic variables and sector effect as well as two levels of random effect with serial dependence. We can tell the difference for companies in different sectors in different year.

4.2.2 Empirical study of rating transition data of n-notches

Data description

We use the Standard & Poor's database CreditPro 6.6 which consists of 10439 companies from 13 industry sectors over the period January 1981 to December 2003. Overall 19054 effective rating migrations are recorded in the CreditPro database. The rating category "NR" is treated the same as in chapter 2. Three companies have starting time after the ending time which is certainly wrong, it might because of errors made by CreditPro database. Since we couldn't get enough information about what had happened, therefore those three companies were removed from our database.

Sector	Name	#observations
1	“Aerospace/automotive/capital goods/metal”	1676
2	“Consumer/service sector”+“Transportation”	2699
3	“Energy and natural resources”+“Utility”	2200
4	“Financial Institutions”+“Insurance”+“Real estate”	3126
5	“Forest and building products”+“Homebuilders”	555
6	“Health care”+ “Chemicals”	859
7	“Hightec/computers/office equipment”+“Telecom”	1235
8	“Leisure time/media”	1176

Table 4.1: Industry sectors used in our analysis

The start rating classes included in our analysis are

$$\mathcal{K} = \{CCC, B, BB, BBB, A, AA, AAA\}$$

We treat default rating D as an absorb state, so the end rating would be include an additional default rating.

$$\mathcal{K}_0 = \{CCC, B, BB, BBB, A, AA, AAA, D\} \cup \{D\}$$

where we merge actual rating k^+, k, k^- into k. We also merge CCC, CC, and C into a single rating class CCC. The whole database consists of 13 industries and 93 countries or region. We use a subset of the data which consists of 6732 US companies with totally 13526 effective rating migrations recorded. The industry sectors in our analysis are different from the S&P industry sectors. We merged 13 industries to 8, Consumer/service with Transportation, Energy with Utility, Financial institution with Insurance and Real estate, Forest and Building products with Homebuilders, Health care with Chemicals and finally High technology with Telecommunications. The industry sectors we merged are have similar business. The details about the industries for the data we used are shown in Table 4.1.

The subset database does not record all the possible rating transitions. We find in total 11 different types of rating transitions as measured by the number of notches. The number of levels (notches) rating transition data we used are shown in Table 4.2.

The multi-state feature of the model is represented as a set of $\mathbb{U}$ of transition types, $\mathbb{U} = \{1, 2, \ldots, U\}$. Standard & Poor's dataset has rating classes $\mathcal{K}_0 = \{CCC, B, BB, BBB, A, A$ $\{D\}$, default rating category D is absorb state. If transition from rating category A to AA is 1 notch upgrade, then we total get {1,2,3,4,5,6} notches upgrade and {1,2,3,4,5,6,7} notches downgrade. Furthermore, there is no observation for 6 and 7 notches downgrade. In Table 4.2 we find that only very few observation available for transition more than 3 notches, therefore we combine all the notches equal or greater than 3 into one transition type. Finally the transition type we considered in this section are $\{1, 2, 3^+\}$ notches upgrade and downgrade, totally 6 transition types so that $\mathbb{U} = \{1, 2, \ldots, 6\}$.

Results

There are many ways to implement Gibbs simulation, but we choose WinBUGS in our analysis as the way to implement. Using R package "R2WinGBUS", WinBugs can be run from R, this package will save your time from click and set up WinBUGS parameters by writing a Bugs model including everything. For running Gibbs simulation in WinBUGS, how to choose the prior specifications are very crucial, we set them up as follows: for all the constant baseline, time-dependent fixed effect macroeconomic variables and time-independent fixed effect sector, the precision τ was set to 0.001, resulting a normal distribution which is very uninformative. A non-informative Gamma prior is assumed for τ, the precision of the frailty parameters. Note that the above 'additive' formulation of the frailty model is equivalent to assuming multiplicative frailties with a log-Normal population distribution.

Model 4A.1: Model with macroeconomic covariates only

The rating migration depend on the "state-of-the-economy". There are several possible proxies and we have shown that Chicago Federal National Activities Index three month moving average (CFNAIMA3) is the best macroeconomic covariate to explain the US economy among them. Therefore we continue to use the CFNAIMA3 as the macroeconomic covariate in our analysis.

With 50000 iterations and 25000 burn in, therefore the first 25000 iterations were discarded. The posterior estimation for Model 4A.1 are given in Table 4.3[1]. The posterior mean of macroeconomic covariate CFNAIMA3 are η and its standard deviation(SD). In our model, η_{d3+} to η_{d1} represents downgrades 3 or more notches to downgrades 1 notch and η_{u1} to η_{u3+} represents upgrades 1 notch to upgrades 3 or more notches. The negative value of η_d and positive value of η_u means the increase of CFNAIMA3 will reduce the intensity rating transitions of downgrade and increase the intensity rating transitions of upgrades. The bigger absolute value of η_d means rating agency much easier to downgrade obliger than upgrade, this could be easily found in the CreditPro database. The posterior mean of υ, which is the parameter for baseline for different rating transitions, $\exp(\upsilon)$ is the baseline in our model for different rating transition type. The value of υ for both downgrade and upgrade are very close which means for the same number of notches, monthly intensity are effected by macroeconomic covariates only. In this simple model, we didn't include yearly shared frailty and will include it in next model. The MC error number is quite small which means the model fits the data well.

Figure 4.1 shows monthly intensity for two notch downgrade and two notch upgrade, the blue line for downgrade and red for upgrade which shows opposite direction. The increasing of macroeconomic covariates will increase the upgrade intensity and decrease the downgrade intensity. Figure 4.2 shows the posterior mean of monthly intensity with 95% credible intervals. The value of η_{d2} η_{u2} are 0.1794 and -0.3680 gives more flat plot for upgrade intensity.

[1] d_n and u_n represent downgrade and downgrade n-notches

Downgrade Notches	5	4	3	2	1	Censor
#Observation	6	27	107	584	4612	5997
Upgrade Notches	1	2	3	4	5	6
#Observation	1982	150	45	14	1	1

Table 4.2: Numbers of notches for Standard & Poors' CreditPro 6.6 from 31/12/1980 - 31/12/2003

Model 4A.1: Model with macroeconomic covariates only						
	η_{d3+}	η_{d2}	η_{d1}	η_{u1}	η_{u2}	η_{u3+}
mean	-0.4292	-0.3680	-0.3367	0.0008	0.1794	0.0361
sd	(0.11)	(0.05268)	(0.0194)	(0.0330)	(0.1289)	(0.1810)
MCerror	0.0007	0.0004	0.0001	0.0002	0.0008	0.0011
Remaining Parameters						
	υ_{d3+}	υ_{d2}	υ_{d1}	υ_{u1}	υ_{u2}	υ_{u3}
mean	-4.271	-3.868	-3.742	-3.769	-3.878	-3.612
sd	0.08813	0.04261	0.01532	0.0224	0.08238	0.1304
MCerror	5.97E-4	3.093E-4	1.132E-4	1.537E-4	4.857E-4	8.706E-4

Table 4.3: Model 4A.1 with macroeconomic variable only

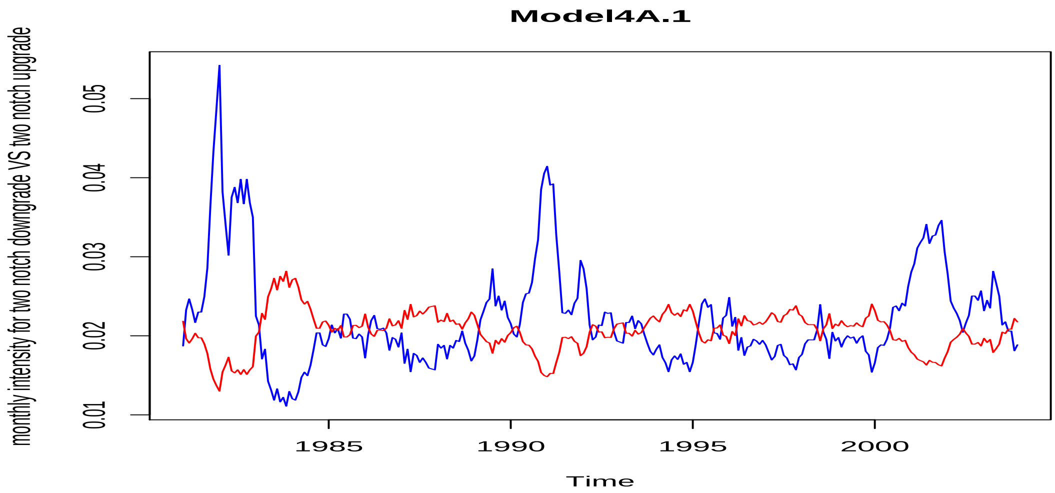

Figure 4.1: Monthly intensity for two notch downgrade VS two notch upgrade, the blue line shows downgrade and red line for upgrade

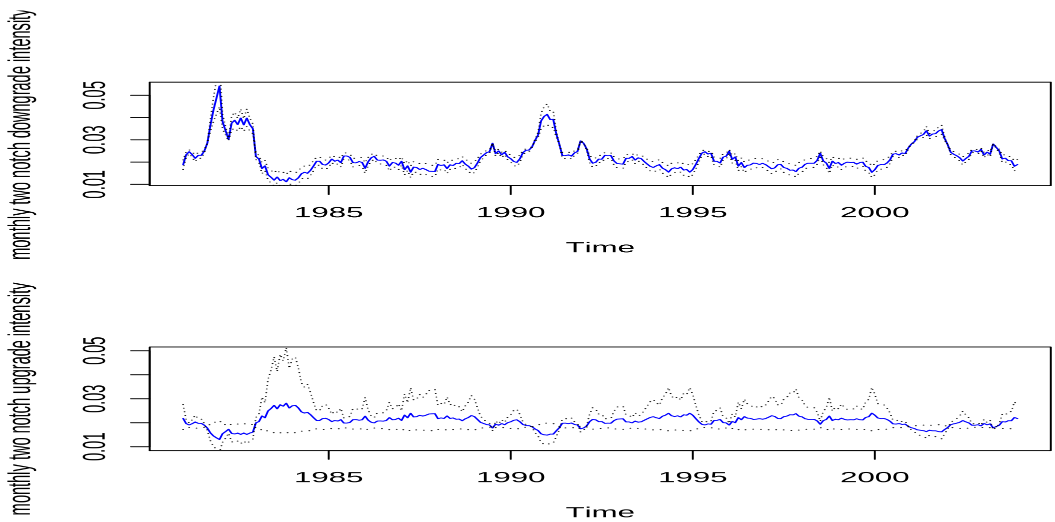

Figure 4.2: Monthly intensity for two notches downgrade and upgrade; posterior means with 95% credible intervals, the upper plot shows downgrade and lower plot for upgrade

Model 4A.2: Model with yearly time-dependent shared frailty

	η_{d3+}	η_{d2}	η_{d1}	η_{u1}	η_{u2}	η_{u3+}	σ
$mean$	-0.2217	-0.2141	-0.1393	0.1551	0.4832	0.1536	0.07915
sd	0.1358	0.06902	0.03681	0.0432	0.1552	0.2021	0.02655
$MCerror$	0.0012	0.0007	0.0005	0.00045	0.0013	0.0014	0.0017

Remaining Parameters

	υ_{d3+}	υ_{d2}	υ_{d1}	υ_{u1}	υ_{u2}	υ_{u3}
$mean$	-4.277	-3.825	-3.785	-3.691	-3.755	-3.532
sd	0.2128	0.1696	0.1515	0.1228	0.212	0.1789
$MCerror$	0.01335	0.0114	0.01049	0.008399	0.01348	0.008138

Table 4.4: Model 4A.2 with yearly time-dependent shared frailty

Model 4A.2: Model with yearly time-dependent shared frailty

The rating migration depend on the "state-of-the-economy". Yearly shared frailty are additional unobserved factors help to explain the rating migration activity, therefore further frailty is needed to capture patterns of variability in responses that cannot be explained by the observed macroeconomic covariates alone. In Model 4A.2, we add yearly shared frailty and try to capture the variability between different years.

As in Model 4A.1, we run 50000 iterations and 25000 burn in, therefore the first 25000 iterations were discarded. We show part of the results for Model 4A.2 in Table 4.4. The macroeconomic covariates have negative effect for downgrade events and positive effect for upgrade event. The MC error number for all parameters are quite small which means the model fits the data well and results are acceptable. The systematic risk factors in Model 4A.1 can be divided into observed fixed effects and unobserved random effect to capture heterogeneity for intensity function. In the following figure, we can see the two risk factors in model 4A.1&2 shoes co-movement.

Figure 4.4 shows monthly intensity for one notch downgrade, the blue line for downgrade with Model4A.2 and red line for Model4A.1. Two lines show some co-movement.

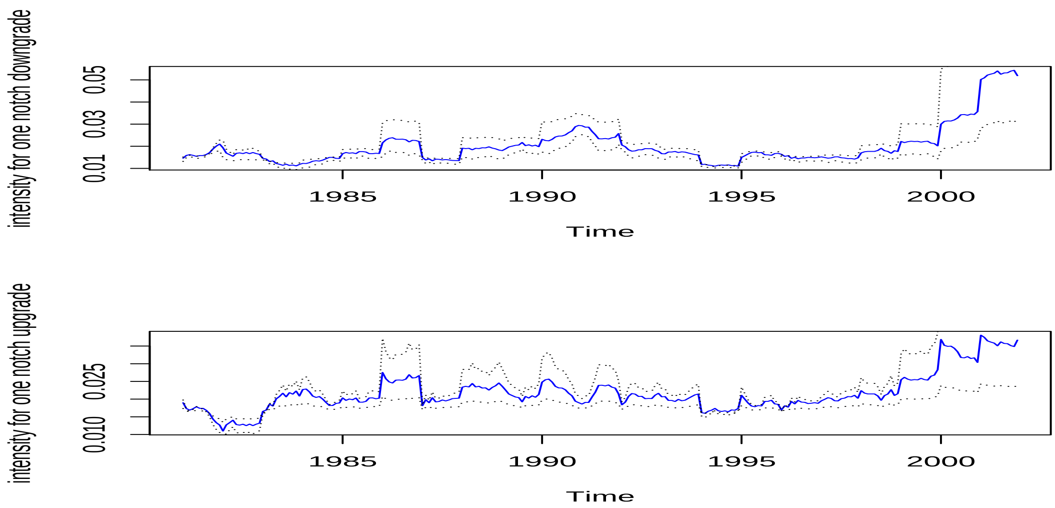

Figure 4.3: Monthly intensity for one notch downgrade and upgrade; posterior means with 95% credible intervals, the upper plot shows downgrade and lower plot for upgrade

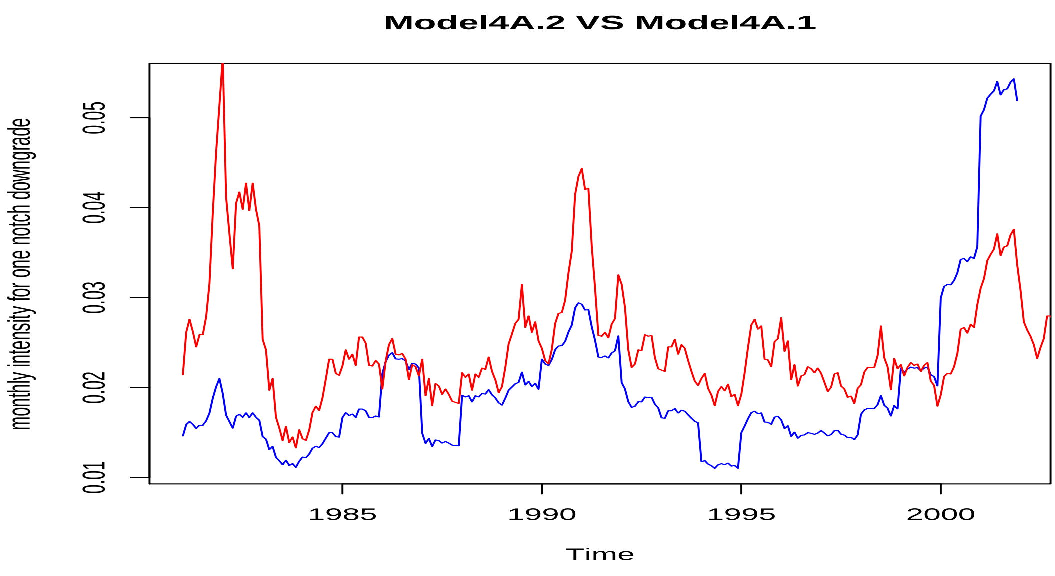

Figure 4.4: Monthly intensity for one notch downgrade for Model 4A.1 (red line) and Model 4A.2 (blue line)

Model 4A.3: Model with serial dependence for yearly shared frailty							
	η_{d3+}	η_{d2}	η_{d1}	η_{u1}	η_{u2}	η_{u3+}	σ
$mean$	-0.2195	-0.2138	-0.1375	0.1546	0.4812	0.1588	0.255
sd	0.1374	0.06785	0.03564	0.04247	0.154	0.2012	0.1042
$MCerror$	0.0013	0.00069	0.00050	0.00045	0.00182	0.00186	0.00734
Remaining Parameters							
	υ_{d3+}	υ_{d2}	υ_{d1}	υ_{u1}	υ_{u2}	υ_{u3}	ϕ
$mean$	-3.196	-2.901	-2.926	-3.0	-2.658	-2.896	0.9361
sd	1.556	1.321	1.232	0.9912	1.609	0.9813	0.05978
$MCerror$	0.123	0.1049	0.09804	0.07877	0.1267	0.0725	0.001816

Table 4.5: Model 4A.3 with serial dependence for yearly shared frailty

Figure 4.3 shows the posterior mean of monthly intensity with 95% credible intervals.

Model 4A.3: Model with serial dependence for yearly shared frailty

The rating migration depend on the “state-of-the-economy”. Frailties are additional unobserved factors help to explain the rating migration activity, therefore further shared frailties are needed to capture patterns of variability in responses that cannot be explained by the observed macroeconomic covariates alone. In Model 4A.3, we add yearly random effect and try to capture the variability between different years.

With 50000 iterations and 25000 burn in, part of the results for Model 4A.3 are presented in Table 4.5. The macroeconomic covariates have negative effect for downgrade events and positive effect for upgrade event. The MC error number for all parameters are quite small which means the model fits the data well and results are acceptable. The systematic risk factors in Model 4A.1 can be divided into observed fixed effects and unobserved random effects to capture heterogeneity for intensity function.In the following figure, we can see the two risk factors in model 4A.1,2&3 shoes co-movement.

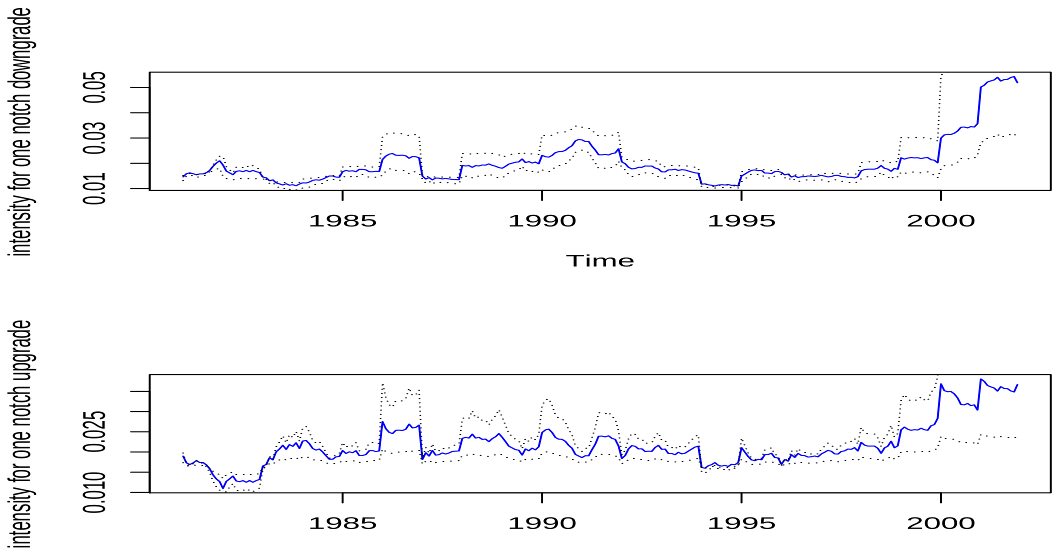

Figure 4.5: Monthly intensity for one notch downgrade and upgrade; posterior means with 95% credible intervals, the upper plot shows downgrade and lower plot for upgrade

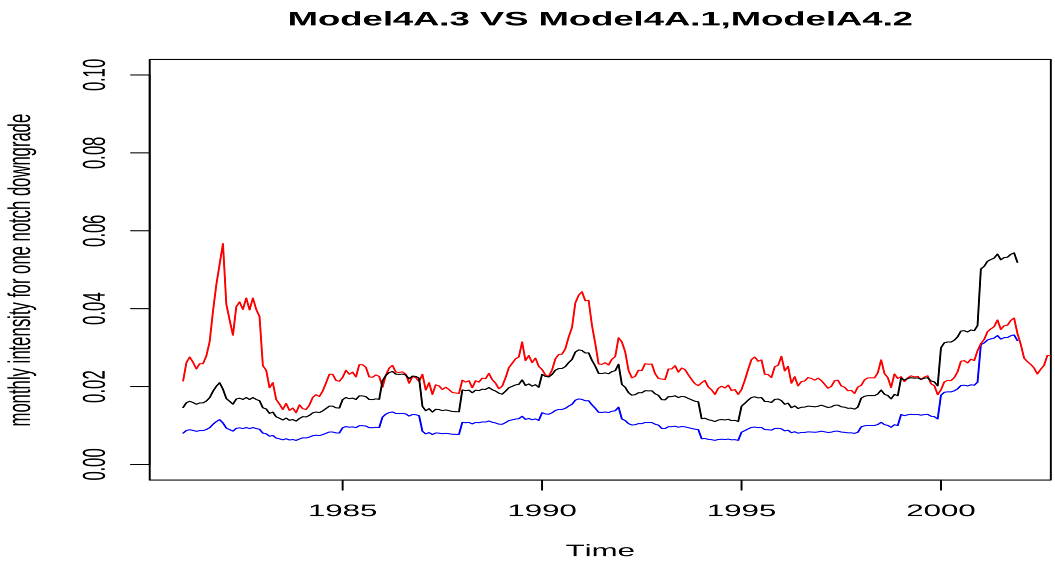

Figure 4.6: Monthly intensity for one notch downgrade for Model 4A.1 (red line), Model 4A.2 (black line) and Model 4A.3 (blue line)

Model 4A.4: Model with shared sector frailties								
	η_{d3+}	η_{d2}	η_{d1}	η_{u1}	η_{u2}	η_{u3+}	σ	σ_1
mean	-0.1892	-0.1876	-0.1309	0.1256	0.4917	0.1389	0.3563	0.2769
sd	0.1397	0.0687	0.0373	0.0419	0.1561	0.1955	0.0824	0.050
MCerror	0.0038	0.0019	0.0014	0.0012	0.0048	0.0039	0.0061	0.0051

Remaining Parameters							
	υ_{d3+}	υ_{d2}	υ_{d1}	υ_{u1}	υ_{u2}	υ_{u3}	ϕ
mean	-2.1916	-2.029	-2.112	-2.512	-1.678	-2.334	0.9482
sd	0.4279	0.3348	0.2905	0.2227	0.4625	0.5331	0.033
MCerror	0.04445	0.03749	0.03385	0.02526	0.04753	0.04181	0.00094

Table 4.6: Model 4A.4: Model with shared sector frailties (results have been omitted for a simpler presentation)

Figure 4.6 shows monthly intensity for one notch downgrade, the blue line for downgrade with Model4A.3 and red line for Model4A.1 and black line for Model4A.3. Three lines show some co-movement. Figure 4.5 shows the posterior mean of monthly intensity with 95% credible intervals.

Model 4A.4: Model with shared sector frailties

We use yearly shared frailty to capture the variability for different time period for any company. Further frailties like industry sectors can be introduced to capture additional variability. We divided the CreditPro data into 8 industry sectors. The following table shows posterior mean for this two-level random effect model.

We run 50000 iterations and discard first 25000 iterations. The part of results for Model 4A.4 are presented in Table 4.6. The macroeconomic covariates have negative effect for downgrade events and positive effect for upgrade event. The MC error number for all parameters are quite small which means the model fits the data well and results are acceptable. The posterior mean for autoregressive coefficient ϕ is

	Dbar	Dhat	pD	DIC
Model 4A.1	1354910.0	1354900.0	11.970	1354920.0
Model 4A.2	1352500.0	1352510.0	-4.462	1352500.0
Model 4A.3	1352500.0	1354160.0	-1656.390	1350850.0
Model 4A.4	1352170.0	1352030.0	133.898	1352300.0

Table 4.7: Deviance information criterion for four different frailty models

0.9482 with standard deviation 0.033. We can easily find the co-movement for each sector as well as the heterogeneity. With introducing the second-level sector random effect, it capture the heterogeneity which cannot be explained by the one-level random effect model. Please see the figure.

Figure 4.7 shows monthly intensity for one notch downgrade with different sectors, these sectors shows co-movement and heterogeneity. Figure 4.8 shows the posterior mean of monthly intensity with 95% credible intervals for sector Finance and Forest. For simplicity, we only pick up these two sectors to see the difference. We can clearly see difference between these two sectors, other sectors also have heterogeneity.

We use deviance information criterion (DIC) to compare these different models.

For each model, the Gibbs sample was run for 50000 iterations, the first 25000 iterations were discarded, and the remaining 25000 iterations for each chain were used for analysis. The estimates for DIC and p_D for the four different models are presented in Table 4.7. The use of DIC is rather to compare different models than choose the true model. Model 4A.2 and Model 4A.3 have been introduced one level shared frailties and there are better than Model 4A.1. And Model 4A.3 has been introduced an AR(1) process, it is better than Model 4A.2 from the DIC estimation. By introducing second level of shared frailties, Model 4A.4 improved compared with Model 4A.2 but do not improved by Model 4A.3 according DIC estimation.

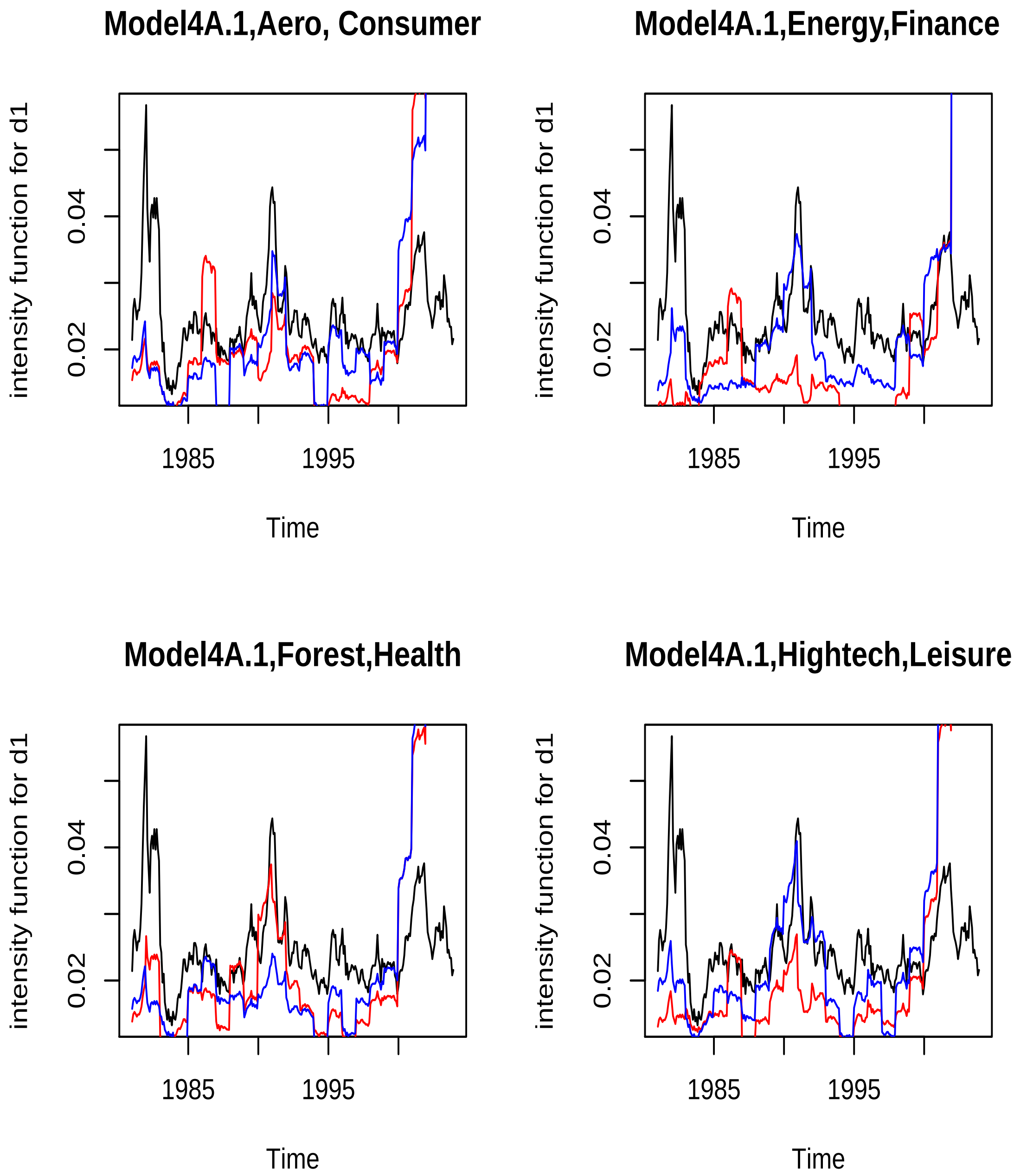

Figure 4.7: Monthly intensity for one notch downgrade with Model 4A.1 (black line), Model 4A.4 (red and blue line) for different sectors

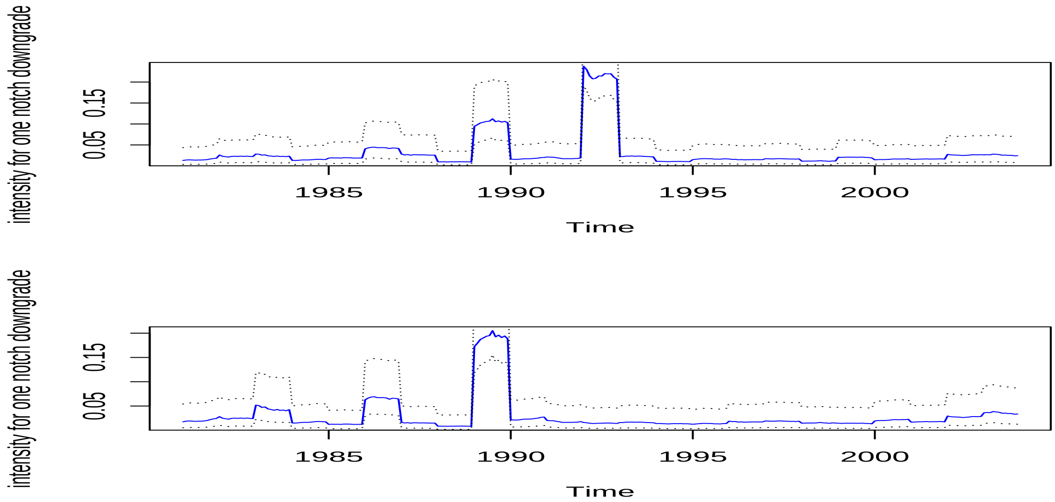

Figure 4.8: Monthly intensity for one notch downgrade; posterior means with 95% credible intervals, the upper plot shows sector Finance and lower plot for sector Forest

4.3 Frailty model for actual credit rating transitions

With the strong assumption that the rating transition of n-notches is the same kind of event, we have merged many different rating transitions into one category in the previous section. However, rating transition of n-notches with different starting rating are different rating transitions. As we know, transition matrices are at the center of modern credit risk management. Transition matrices are widely used for risk management purpose, economic capital purpose and credit derivatives pricing. But we still cannot get estimated transition matrices with the previous model, therefore we will extend the previous simple model to allow for all the actual rating transitions which can used for transition matrices calculation.

4.3.1 Theoretical models for actual credit rating transition

Model 4B.1: Model with macroeconomic covariates only

Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}_t$ and actual rating transition effect υ_{hj} for rating transition (h, j). The monthly intensity function is given by:

$$\lambda_{hji}(t|\boldsymbol{\eta}, \upsilon) = Y_i(t)\exp(\boldsymbol{\eta}'_{hj}\boldsymbol{z}(t) + \upsilon_{hj}) \tag{4.8}$$

This model only depend on macroeconomic covariates for each transition (h, j), frailty is not considered in this simple model.

Model 4B.2: Model with yearly time-dependent shared frailty

The model with macroeconomic covariates will be extended by adding a component of yearly heterogeneity $b_{y(t)}$, this frailty process is only depend on time t for each transition (h, j) and will take the number yearly. Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}_t$, rating transition effect $\boldsymbol{\upsilon}$ and yearly shared frailty. The monthly intensity function is given by:

$$\lambda_{hji}(t|\boldsymbol{\eta}, \upsilon, \boldsymbol{\gamma}) = Y_i(t)\exp(\boldsymbol{\eta}'_{hj}\boldsymbol{z}_(t) + \upsilon_{hj} + \boldsymbol{\gamma}_{hj}b_{y(t)}) \tag{4.9}$$

where $y(t)$ gives the year period corresponding to time t. For each rating transition (h, j), macroeconomic covariates have different parameter $\boldsymbol{\eta}$ and different frailty process b_t. This model allows for difference in the events times that depend on the macroeconomic variables, rating transition effect as well as the frailty for each year. However, it does not allow serial dependence between each yearly shared frailty for different rating transition.

Model 4B.3: Model with serial dependence for yearly shared frailty

The model with yearly shared frailty will be extended by adding serial dependence for yearly heterogeneity $b_{y(t)}$. Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}_t$, rating transition effect υ_{hj} and yearly shared frailty. The monthly intensity function is given by:

$$\lambda_{hji}(t|\boldsymbol{\eta}, \upsilon, \boldsymbol{\gamma}) = Y_i(t) \exp(\boldsymbol{\eta}'_{hj} \boldsymbol{z}(t) + \upsilon_{hj} + \boldsymbol{\gamma}_{hj} b_{y(t)}) \tag{4.10}$$

$$b_y = \varphi b_{y-1} + \epsilon_y$$

A simple model AR(1) process has been used to capture the serial dependence for yearly shared frailty. This model allows for difference in the events times that depend on the macroeconomic variables, rating transition effect as well as yearly shared frailty with serial dependence for different rating transition.

Model 4B.4: Model with shared sector frailties

We will add shared sector frailty as the second level of frailty. Then the shared frailty with two levels will become $b_{y(t),s(i)}$. The model with yearly shared frailty will be extended by adding serial dependence for yearly heterogeneity $b_{y(t)}$. Suppose the intensity function depends on time-dependent macroeconomic covariates $\boldsymbol{z}_t$, rating transition effect and two levels of shared frailty of yearly and sector. The monthly intensity function is given by:

$$\lambda_{hji}(t|\boldsymbol{\eta}, \upsilon, \boldsymbol{\gamma}) = Y_i(t) \exp(\boldsymbol{\eta}'_{hj} \boldsymbol{z}(t) + \upsilon_{hj} + \boldsymbol{\gamma}_{hj} b_{y(t),s(i)}) \tag{4.11}$$

$$b_{y,s} \sim N(b_y, \sigma^2)$$

$$b_y = \varphi b_{y-1} + \epsilon_y$$

A simple model AR(1) process has been used to capture the serial dependence for yearly shared frailty. This model allows for difference in the events times that depend on the macroeconomic variables and rating transition effect as well as two levels of random effect with serial dependence. We can tell the difference for companies in different sectors in different year.

	AAA	AA	A	BBB	BB	B	CCC	D	Total
AAA	206	148	14	2	2	0	0	0	372
AA	44	481	553	34	5	5	1	0	1123
A	11	285	1246	840	57	27	1	5	2472
BBB	5	27	521	1342	659	72	9	19	2654
BB	3	8	46	483	1147	883	53	64	2687
B	1	7	25	48	520	1382	850	354	3187
CCC	1	0	4	7	18	129	193	679	1031
Total	271	956	2409	2756	2408	2498	1107	1121	13526

Table 4.8: Numbers of transitions for Standard & Poors' CreditPro 6.6 from 31/12/1980 - 31/12/2003

4.3.2 Empirical study of rating transition data

Data description

In the previous section, we study numbers of notches for credit rating transition. In this section, the subset database will record all the possible rating transitions. We totally find 44 different rating transitions and censored database.

The multi-state feature of the model is represented as a set of $\mathbb{U}$ of transition types, $\mathbb{U} = \{1, 2, \ldots, U\}$. Standard & Poor's dataset has rating classes $\mathcal{K}_0 = \{CCC, B, BB, BBB, A, A$ $\{D\}$, default rating category D is absorb state. Therefore the total number of possible rating transition types is 49 and censored case. There are no transitions recorded in Standard & Poor's for rating AAA to B,CCC and D, rating AA to D and CCC to AA. So the total number of transition types which has been considered in our analysis is $S = 44$.

Results

We use "R2WinGBUS" to implement Gibbs simulation for actual rating transition estimation in this section. How to set up prior distribution is very crucial. We use similar model as in the previous section, with the prior specifications as follows: For all the constant baseline, time-dependent fixed effect macroeconomic variables, the precision τ was set to 0.001, resulting a normal distribution which is very uninformative. A non-informative Gamma prior is assumed for τ, the precision of the frailty parameters. Note that the above 'additive' formulation of the frailty model is equivalent to assuming multiplicative frailties with a log-Normal population distribution.

Model 4B.1: Model with macroeconomic covariates

The rating migration depend on the "state-of-the-economy". There are several possible proxies and we have shown that Chicago Federal National Activities Index three month moving average (CFNAIMA3) is the best macroeconomic covariate to explain the US economy among them. Therefore we will use the CFNAIMA3 as the macroeconomic covariates in this analysis.

With 50000 iterations and 25000 burn in, therefore the first 25000 iterations were discarded. In this model, we have totally 44 actual rating transitions. The posterior mean of macroeconomic covariate CFNAIMA3 η and its standard deviation(SD) are shown in Table 4.9. The results show that the increase of CFNAIMA3 will reduce the intensity rating transitions of downgrade and increase the intensity rating transitions of upgrades. Most of the downgrades have negative η and upgrades have positive η. However, some parameters have large errors in such a complicated model especially for the low number transition types. The posterior mean of υ, which is the parameter for baseline for different rating transitions, $\exp(\upsilon)$ is the baseline in our model for different rating transition type. In this simple model, we didn't include frailty and will include it in next model. The MC error number is quite small which means the model fits the data well. The results are omitted from showing in the content.

Model 4B1: Posterior mean for η

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.2420	-0.3544	0.2461	-9.0700	0	0	0
AA	0.0653	-	-0.3217	0.2105	-0.7040	-0.1849	5.6450	0
A	0.3534	-0.0302	-	-0.3150	-0.1059	0.4394	0.6133	-0.5849
BBB	-0.7852	0.5325	0.1004	-	-0.2500	0.0285	-0.4717	-0.4969
BB	0.5113	0.3202	0.2072	-0.1205	-	-0.2534	-0.5838	-0.8407
B	0.3892	-0.3512	0.1565	-0.1103	0.0036	-	-0.6255	-0.5560
CCC	-11.930	0	-6.2420	1.0910	0.5004	0.06764	-	-0.3838

Model 4B1: Posterior standard deviation for η

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.1061	0.3907	1.3900	4.6150	0	0	0
AA	0.2091	-	0.0550	0.2340	0.5027	0.4985	8.0550	0
A	0.3971	0.0785	-	0.0440	0.1740	0.2525	1.2640	0.6212
BBB	0.5775	0.3247	0.0620	-	0.0534	0.1371	0.5295	0.3016
BB	1.2360	0.5314	0.2132	0.0704	-	0.0488	0.1745	0.1633
B	1.1520	0.4961	0.2816	0.2248	0.0709	-	0.0494	0.0703
CCC	10.230	0	2.5040	1.0180	0.4558	0.1320	-	0.0571

Model 4B1: Posterior mean for υ

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	-5.2020	-5.4460	-5.3590	-4.0210	0	0	0
AA	-4.836	-	-4.982	-5.204	-4.774	-4.543	-5.105	0
A	-3.574	-3.984	-	-4.117	-3.895	-4.259	-5.659	-5.153
BBB	-6.477	-5.944	-5.633	-	-5.548	-5.505	-4.982	-6.444
BB	-5.416	-5.398	-5.327	-5.301	-	-5.059	-4.972	-6.049
B	-6.614	-5.21	-5.097	-5.409	-5.090	-	-4.88	-5.263
CCC	-9.343	0	-3.996	-4.876	-5.710	-4.842	-	-4.304

Model 4B1: Posterior standard deviation for υ

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.0844	0.2877	0.8524	1.0740	0	0	0
AA	0.1531	-	0.0437	0.1762	0.7067	0.4924	1.8850	0
A	0.3196	0.06024	-	0.03622	0.132	0.2072	1.544	0.5364
BBB	0.6020	0.2137	0.04307	-	0.0406	0.1200	0.3818	0.2433
BB	0.7145	0.3826	0.1501	0.04526	-	0.03494	0.1523	0.1425
B	1.5590	0.4101	0.2053	0.1470	0.04387	-	0.0373	0.0573
CCC	4.5720	0	0.5961	0.5037	0.2405	0.0897	-	0.04174

Table 4.9: Model 4B.1 with macroeconomic variable

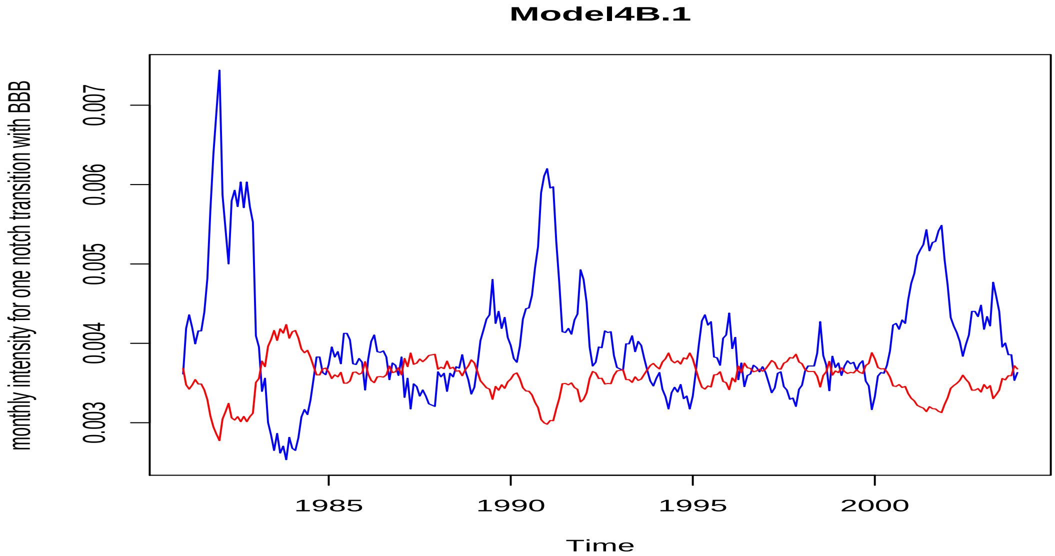

Figure 4.9: Monthly intensity for downgrade and upgrade for rating BBB, the blue line shows downgrade and red line for upgrade

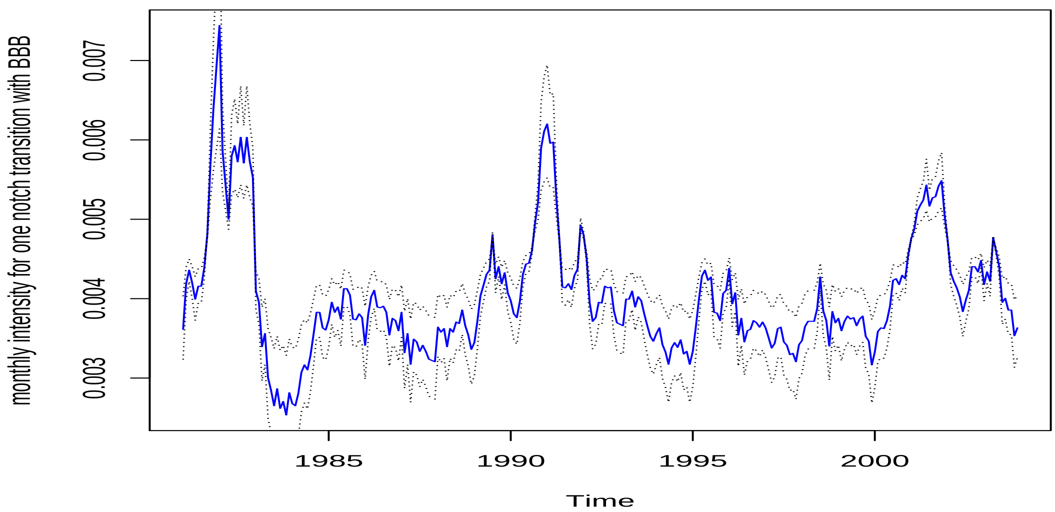

Figure 4.10: Monthly intensity for transition BBB to BB; posterior means with 95% credible intervals

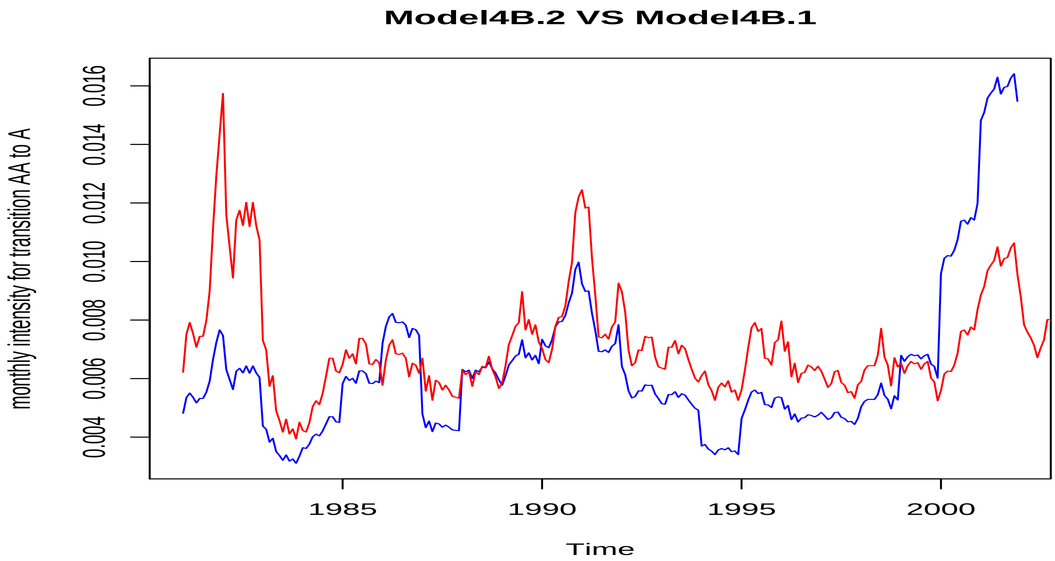

Figure 4.11: Monthly intensity for rating transition AA to A, the blue line shows Model 4B.2 and red line for Model 4B.1

Figure 4.9 shows monthly intensity for rating transition BBB to BB and BBB to A, the downgrade (blue) and upgrade (red) show different effect with macroeconomic covariates. Downgrade and upgrade are showing opposite direction, and downgrade normally has larger intensity than upgrade because rating agency are easier to downgrade than upgrade. Figure 4.10 shows the posterior mean of monthly intensity with 95% credible intervals for rating transition BBB to BB.

Model 4B.2: Model with yearly shared frailty

The rating migration depend on the "state-of-the-economy". Frailties are additional unobserved factors help to explain the rating migration activity, therefore further frailty is needed to capture patterns of variability in responses that cannot be explained by the observed macroeconomic covariates alone. In Model 4B.2, we add yearly shared frailty and try to capture the variability between different years.

As in Model 4B.1, we run 50000 iterations and 25000 burn in, therefore the first 25000 iterations were discarded. We show the results for Model 4B.2 in Table 4.10 and 4.11.

Model 4B2: Posterior mean for η								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.3338	-0.0628	0.6204	-9.0560	0	0	0
AA	0.2244	-	-0.1782	0.4820	-0.8058	-0.0789	5.8960	0
A	0.3784	0.0318	-	-0.1784	0.0479	0.6111	0.6841	-0.5099
BBB	-0.7863	0.7776	0.2434	-	-0.0799	0.1129	-0.3425	-0.3005
BB	1.3750	0.4254	0.4231	-0.0116	-	-0.0488	-0.5222	-0.6817
B	0.4341	-0.3375	0.2540	0.2102	0.2051	-	-0.4044	-0.4265
CCC	-11.2800	0	-6.2770	1.6840	0.7777	0.3400	-	-0.2935

Model 4B2: Posterior standard deviation for η								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.1190	0.4811	1.6260	4.6060	0	0	0
AA	0.2362	-	0.0746	0.2792	0.5474	0.5786	8.4540	0
A	0.4133	0.08767	-	0.06279	0.2141	0.3212	1.326	0.6877
BBB	0.6387	0.3671	0.0757	-	0.0672	0.1555	0.5823	0.3736
BB	1.6230	0.5917	0.2452	0.0795	-	0.0636	0.1956	0.2011
B	1.1490	0.5140	0.3215	0.2720	0.0829	-	0.0647	0.0850
CCC	10.440	0	2.4610	1.2770	0.5232	0.1584	-	0.0641

Model 4B2: Posterior mean for υ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	-5.120	-5.596	-5.650	-4.018	0	0	0
AA	-4.772	-	-4.953	-5.148	-4.900	-4.232	-5.141	0
A	-3.551	-3.876	-	-4.106	-3.772	-3.977	-5.616	-5.297
BBB	-6.627	-5.79	-5.536	-	-5.630	-5.405	-5.108	-6.527
BB	-5.708	-5.462	-5.220	-5.242	-	-5.156	-5.018	-6.231
B	-6.524	-5.084	-4.877	-5.401	-5.110	-	-5.041	-5.287
CCC	-9.033	0	-4.063	-5.080	-5.682	-4.965	-	-4.423

Model 4B2: Posterior standard deviation for υ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.1573	0.3471	0.9940	1.1080	0	0	0
AA	0.2358	-	0.2241	0.3085	0.7720	0.6656	2	0
A	0.4109	0.1548	-	0.2045	0.3174	0.4412	1.6280	0.6029
BBB	0.6733	0.3278	0.1788	-	0.1627	0.2087	0.4348	0.3704
BB	0.8403	0.4155	0.2719	0.1243	-	0.1416	0.2246	0.2868
B	1.5860	0.5105	0.3767	0.2686	0.1311	-	0.1606	0.1887
CCC	4.6990	0	0.6746	0.5924	0.3142	0.1597	-	0.0853

Table 4.10: Model 4B.2 with yearly shared frailties

Model 4B2: Posterior mean for γ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	16.210	20.130	13.530	3.8680	0	0	0
AA	21.490	-	26.690	29.740	14.450	37.800	3.010	0
A	8.844	17.120	-	24.720	34.120	37.60	10.360	18.080
BBB	15.830	28.710	21.070	-	19.140	20.760	9.576	30.220
BB	19.550	10.360	27.030	14.000	-	16.790	19.380	28.810
B	8.100	22.250	28.490	27.220	14.950	-	19.160	21.950
CCC	4.155	0	15.340	16.590	21.200	15.860	-	8.696
Model 4B2: Posterior standard deviation for γ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	4.9650	7.4000	22.140	31.670	0	0	0
AA	6.6430	-	3.9470	7.0060	18.840	28.560	32.030	0
A	23.930	3.9130	-	3.5790	7.6270	23.330	31.700	12.130
BBB	14.190	12.090	3.5060	-	2.8990	5.4050	10.810	7.962
BB	20.120	13.840	8.9910	2.9500	-	2.4940	5.4160	5.2680
B	31.110	27.450	19.560	6.4060	2.5470	-	2.8460	3.5460
CCC	31.350	0	31.070	17.620	11.210	3.0370	-	1.5780

Table 4.11: Model 4B.2 with yearly shared frailties

Compare to Model 4B.1, we add yearly shared frailty ($\gamma_{hj}b_{y(t)}$) into the model. The systematic risk factors in Model 4B.1 can be divided into observed fixed effects and unobserved yearly shared frailty to capture heterogeneity for intensity function. The macroeconomic covariates have negative effect for downgrade events and positive effect for upgrade event. Most of the results are acceptable especially near the diagonal but some off-diagonal η with low transitions are not reasonable. The posterior mean of υ, which is the parameter for baseline for different rating transitions, $\exp(\upsilon)$ is the baseline in our model for different rating transition type. γ is the parameter for yearly shared frailty. The MC error number for all parameters are quite small which means the model fits the data well and results are acceptable. Figure 4.11 shows monthly intensity for rating transition AA to A, Model 4B.1 and Model 4B.2, we can see the two risk factors in Model 4B.1&2 shows co-movement. Figure 4.12 shows the posterior mean of monthly intensity with 95% credible intervals for rating transition AA to A.

Figure 4.13 compares macroeconomic covariates and yearly shared frailties $\boldsymbol{\gamma}_{AA\rightarrow A}b_{y(t)}$. Both γ and b_t are calibrated from model, therefore these two parameters are effected

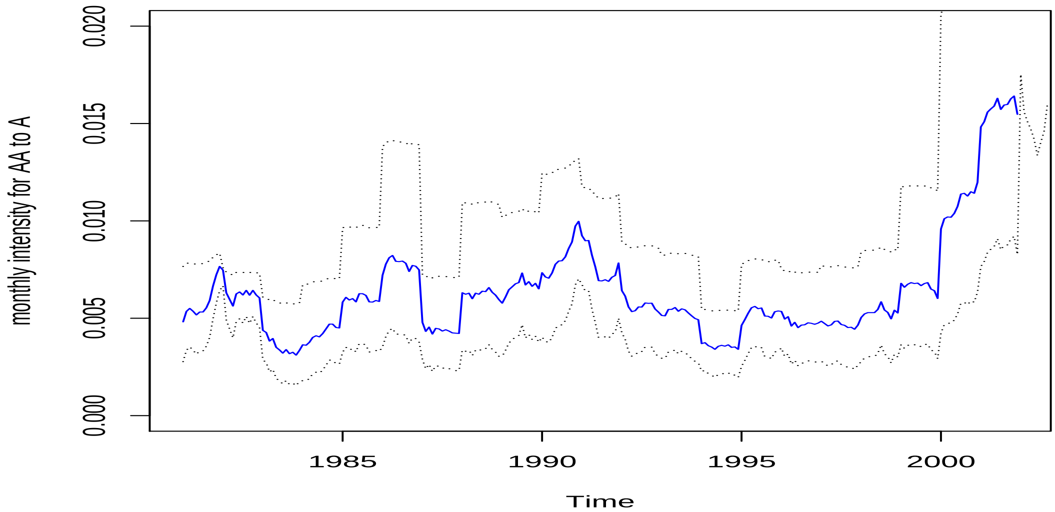

Figure 4.12: Monthly intensity for transition AA to A; posterior means with 95% credible intervals

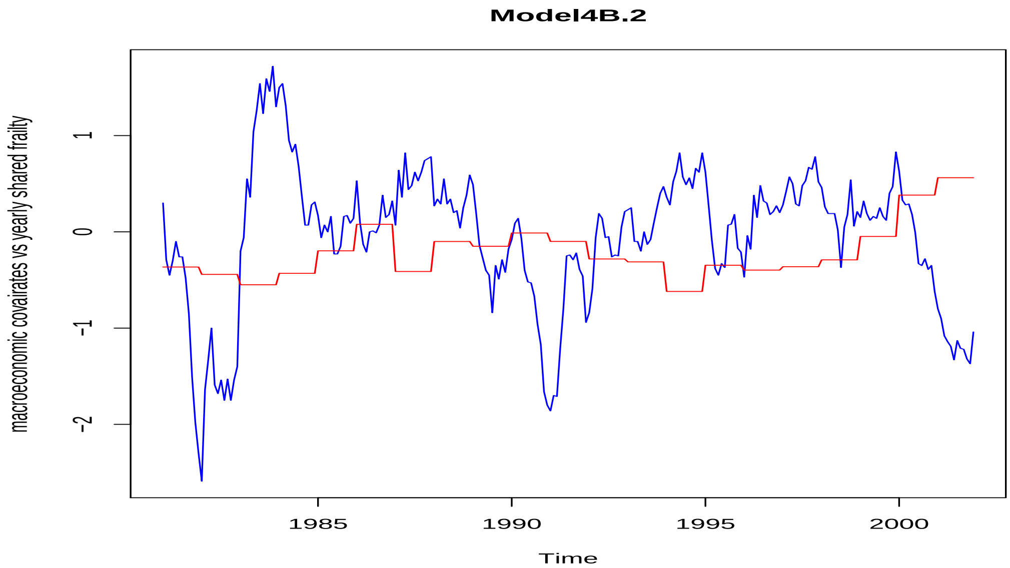

Figure 4.13: Macroeconomic covariates vs yearly shared frailty

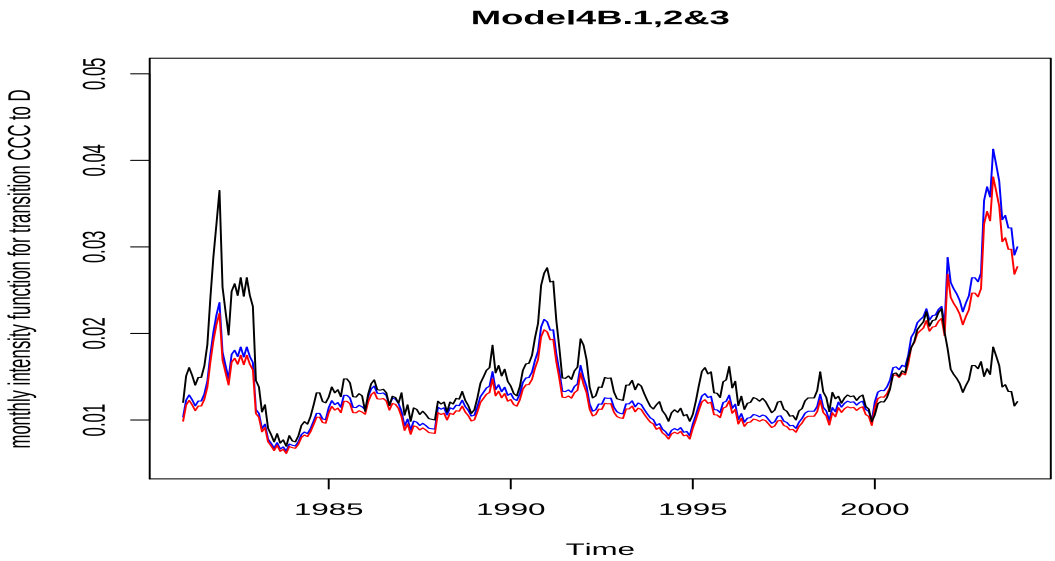

Figure 4.14: Monthly intensity for rating transition CCC to D, the blue line shows Model 4B.3, red line for Model 4B.2 and black for Model 4B.1

by each other. It is better to compare macroeconomic covariates with $\boldsymbol{\gamma}_{AA\to A}b_{y(t)}$. $\eta_{AA\to A}z_t$ and$\boldsymbol{\gamma}_{AA\to A}b_{y(t)}$ will have co-movement.

Model 4B.3: Model with serial dependence for yearly shared frailty

The rating migration depend on the "state-of-the-economy". Frailties are additional unobserved factors help to explain the rating migration activity, therefore further shared frailty is needed to capture patterns of variability in responses that cannot be explained by the observed macroeconomic covariates alone. In Model 4B.3, we add yearly shared frailty with serial dependence and try to capture the variability between different years.

With 50000 iterations and 25000 burn in, part of the results for Model 4B.3 are presented in Table 4.12 and 4.13. Compare to Model 4B.2, we add serial dependence for yearly shared frailty ($\gamma_{hj}b_{y(t)}$). The systematic risk factors in Model 4B.1 can be divided into observed fixed effects and unobserved yearly shared frailty to capture heterogeneity for intensity function. The macroeconomic covariates have negative

Model 4B3: Posterior mean for η								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.3370	-0.02780	0.9224	-9.2150	0	0	0
AA	0.2283	-	-0.1723	0.5078	-0.8370	0.0398	5.9900	0
A	0.3742	0.0351	-	-0.1759	0.0582	0.7061	0.7204	-0.5090
BBB	-0.7951	0.8226	0.2483	-	-0.0802	0.1243	-0.3475	-0.2888
BB	1.7950	0.4410	0.4348	-0.0091	-	-0.0474	-0.5194	-0.6834
B	0.4454	-0.3506	0.2947	0.2230	0.2056	-	-0.4024	-0.4269
CCC	-11.950	0	-6.2490	1.8680	0.7963	0.3397	-	-0.2898

Model 4B3: Posterior standard deviation for η								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.1187	0.4841	1.8340	4.7910	0	0	0
AA	0.2383	-	0.0745	0.2913	0.5566	0.6968	8.5190	0
A	0.4176	0.0860	-	0.0614	0.2120	0.3512	1.3830	0.6969
BBB	0.6428	0.3822	0.0733	-	0.0676	0.1600	0.5932	0.3797
BB	1.8710	0.5902	0.2462	0.0789	-	0.0628	0.1931	0.1999
B	1.1790	0.5439	0.3244	0.2727	0.0827	-	0.0640	0.0844
CCC	10.550	0	2.5530	1.3550	0.5325	0.1568	-	0.06442

Model 4B3: Posterior mean for υ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	-4.923	-5.339	-5.688	-3.780	0	0	0
AA	-4.506	-	-4.612	-4.753	-4.736	-3.023	-4.907	0
A	-3.382	-3.658	-	-3.793	-3.331	-3.308	-5.091	-5.076
BBB	-6.427	-5.421	-5.270	-	-5.388	-5.134	-4.999	-6.143
BB	-5.543	-5.331	-4.877	-5.070	-	-4.945	-4.778	-5.87
B	-5.937	-4.412	-4.350	-5.057	-4.922	-	-4.798	-5.01
CCC	-9.157	0	-4.769	-4.916	-5.399	-4.764	-	-4.314

Model 4B3: Posterior standard deviation for υ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.3218	0.4995	1.1960	1.7240	0	0	0
AA	0.4446	-	0.4934	0.5925	0.9000	1.8940	2.3550	0
A	0.8029	0.3264	-	0.4568	0.6618	1.1530	2.2710	0.6793
BBB	0.7372	0.6745	0.3898	-	0.3530	0.4179	0.4862	0.6225
BB	1.0160	0.5345	0.5642	0.2640	-	0.3140	0.4015	0.5564
B	2.3510	1.4060	0.9247	0.5436	0.2792	-	0.3552	0.4079
CCC	4.7030	0	1.6740	0.7557	0.5452	0.3106	-	0.1674

Table 4.12: Model 4B.3 with serial dependence for yearly shared frailties

Model 4B3: Posterior mean for γ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	-10.100	-13.330	-13.670	-7.413	0	0	0
AA	-13.830	-	-17.040	-19.400	-11.000	-45.430	-6.775	0
A	-7.283	-10.860	-	-15.760	-22.110	-32.700	-16.770	-12.280
BBB	-11.090	-19.600	-13.400	-	-12.150	-13.450	-6.229	-19.610
BB	-16.690	-6.985	-17.690	-8.875	-	-10.730	-12.440	-18.520
B	-15.090	-27.740	-23.210	-17.700	-9.526	-	-12.240	-13.97
CCC	-6.965	0	-23.370	-12.54	-14.450	-10.160	-	-5.548

Model 4B3: Posterior standard deviation for γ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	3.6410	5.3210	17.430	30.450	0	0	0
AA	4.6090	-	3.4780	5.3840	13.340	25.010	31.610	0
A	17.960	3.0080	-	3.2640	5.7430	18.050	30.170	7.9800
BBB	9.1870	8.1640	3.0140	-	2.5370	4.0020	7.1660	5.7300
BB	15.300	9.3770	6.4930	2.3270	-	2.2560	3.9540	4.3110
B	31.090	24.940	14.000	5.0150	2.1570	-	2.5490	3.1060
CCC	31.960	0	31.540	12.680	7.6600	2.5130	-	1.3030

Table 4.13: Model 4B.3 with serial dependence for yearly shared frailties

effect for downgrade events and positive effect for upgrade event. Most of the results are acceptable especially near the diagonal but some off-diagonal η with low transitions are not reasonable. The posterior mean of υ, which is the parameter for baseline for different rating transitions, $\exp(\upsilon)$ is the baseline in our model for different rating transition type. γ is the parameter for yearly shared frailty. The MC error number for all parameters are quite small which means the model fits the data well and results are acceptable. Figure 4.14 shows monthly intensity for rating transition type CCC to D which is different from figure 4.11. We can see the three risk factors in Model 4B.1,2&3 shoes co-movement, especially for Model 4B.2 and 4B.3. The only difference of these two models are serial dependence. The yearly shared frailty makes it different from Model 4B.1.

Figure 4.15 shows the posterior mean of monthly intensity with 95% credible intervals for rating transition CCC to D with Model 4B.3.

Here we consider rating transition $AA \to A$ again, Figure 4.16 compares macroeconomic covariates and yearly shared frailties $\gamma_{AA\to A} b_{y(t)}$ in Model 4B.2 and 4B.3.

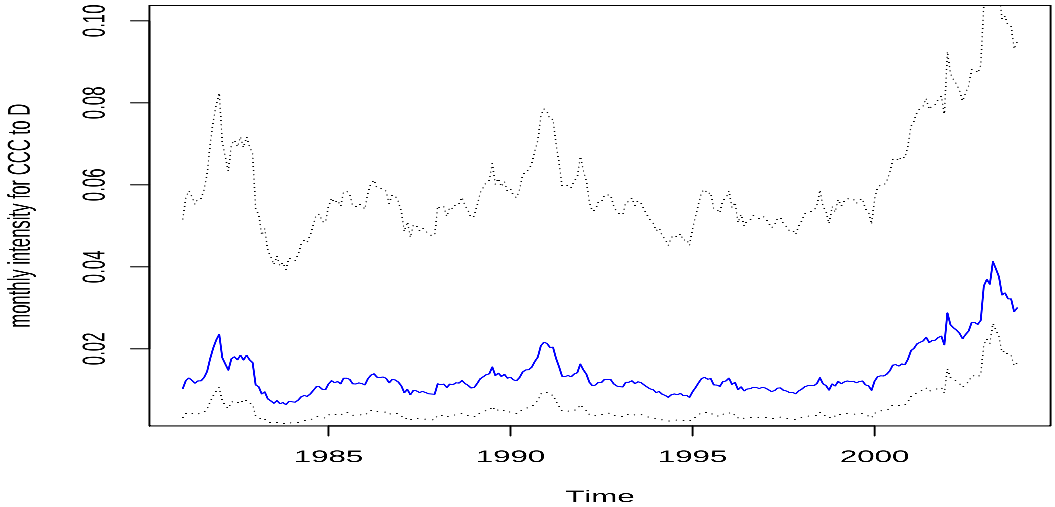

Figure 4.15: Monthly intensity for transition CCC to D; posterior means with 95% credible intervals

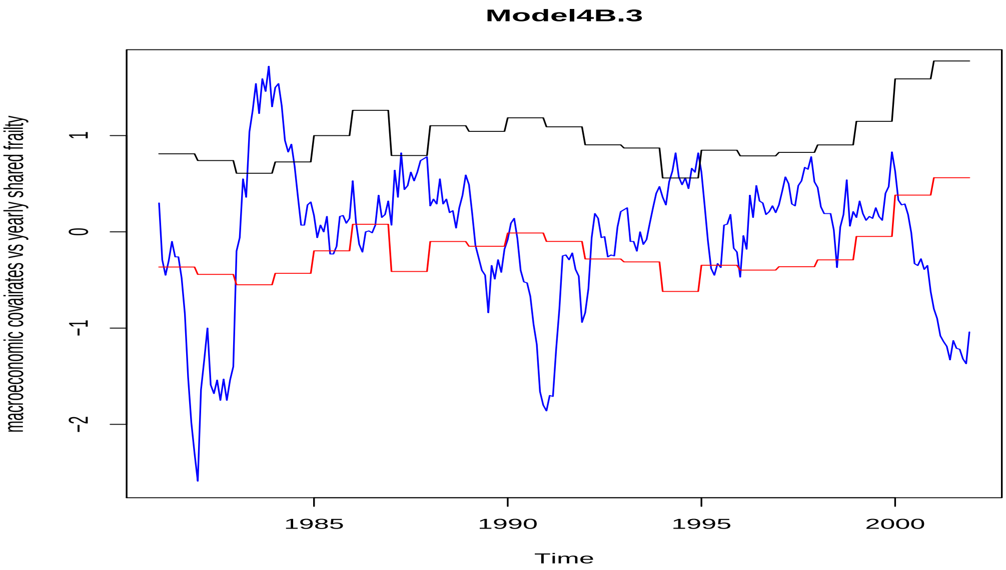

Figure 4.16: Macroeconomic covairates vs yearly shared frailty

$\gamma_{AA\to A} b_{y(t)}$ in Model 4B.2 and 4B.3 are difference, but the difference are from $\upsilon_{AA\to A}$ in two models.

Model 4B.4: Model with shared sector frailties

We use yearly shared frailty to capture the variability for different time period for any company. Further frailties like industry sectors can be introduced to capture additional variability. We divided the CreditPro data into 8 industry sectors. The following table shows posterior mean for this two-level frailties model.

We run 50000 iterations and discard first 25000 iterations. Part of results for Model 4B.4 are presented in Table 4.14 and 4.15. Compare to Model 4B.2, we add second level of shared sector frailty ($\gamma_{hj} b_{y(t)}$). The systematic risk factors in Model 4B.1 can be divided into observed fixed effects and unobserved to capture heterogeneity for intensity function. The macroeconomic covariates have negative η for downgrade events and positive η for upgrade event. Most of the results are acceptable especially near the diagonal but some off-diagonal η with low transitions are not reasonable. $\eta_{A\to AA}$ is negative but very close to zero but $\eta_{AAA\to AA}$ is positive. The posterior mean of υ, which is the parameter for baseline for different rating transitions, $\exp(\upsilon)$ is the baseline in our model for different rating transition type. γ is the parameter for yearly shared frailty. The MC error number for all parameters are quite small which means the model fits the data well and results are acceptable. The MC error number for all parameters are quite small which means the model fits the data well and results are acceptable. The posterior mean for autoregressive coefficient ϕ is 0.9482 with standard deviation 0.033. We can easily find the co-movement for each sector as well as the heterogeneity. With introducing the second-level sector frailties, it capture the sectoral heterogeneity which cannot be explained by the one-level shared frailty model.

Figure 4.17 shows monthly intensity for rating transition BB to B with different sectors, these sectors shows co-movement and heterogeneity. The $\eta_{BB\to B}$ is -0.0025 which is very close to zero, this is why this figure is different from Figure 4.7. In

Model 4B4: Posterior mean for η								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.3408	-0.02390	1.2990	-18.620	0	0	0
AA	0.1815	-	-0.1922	0.5127	-0.8008	0.9868	8.0270	0
A	0.2950	-0.0017	-	-0.1886	0.0533	1.2220	2.2690	-0.5462
BBB	-0.8305	0.8131	0.1898	-	-0.0633	0.1393	-0.6461	-0.2345
BB	2.2490	0.4048	0.4001	-0.0256	-	-0.0025	-0.4369	-0.6017
B	1.4670	-0.2218	0.2177	0.2593	0.2187	-	-0.3230	-0.3797
CCC	-20.130	0	-9.2730	1.6090	1.0310	0.3647	-	-0.2641

Model 4B4: Posterior standard deviation for η								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.1193	0.4897	1.9860	10.010	0	0	0
AA	0.2372	-	0.07181	0.2901	0.5996	1.181	9.9870	0
A	0.4338	0.0849	-	0.0620	0.2086	0.4660	3.2200	0.6855
BBB	0.6699	0.3745	0.0670	-	0.0689	0.1552	0.6095	0.3750
BB	2.0830	0.5879	0.2447	0.0778	-	0.0653	0.2001	0.2081
B	1.9440	0.6155	0.3067	0.2731	0.0835	-	0.0685	0.0889
CCC	14.460	0	4.5990	1.2470	0.5754	0.1605	-	0.0641

Model 4B4: Posterior mean for υ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	-4.6520	-4.8790	-4.6140	5.2750	0	0	0
AA	-4.2820	-	-4.1230	-4.0290	-4.1870	4.4930	-5.3490	0
A	-4.0550	-3.4690	-	-3.2330	-2.7350	-0.7678	3.8440	-4.6180
BBB	-6.0230	-5.0270	-4.9940	-	-4.9070	-4.7650	-5.4190	-5.3250
BB	-4.2070	-5.2560	-4.5190	-4.8580	-	-4.4940	-4.1620	-5.0270
B	2.7030	-2.2090	-4.4800	-4.3690	-4.5660	-	-4.2780	-4.4150
CCC	-8.7840	0	-13.680	-4.7130	-4.6780	-4.3580	-	-4.0880

Model 4B4: Posterior standard deviation for υ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.3472	0.4899	1.7340	7.7940	0	0	0
AA	0.3734	-	0.4509	0.6014	1.0160	4.8730	12.840	0
A	1.1660	0.2685	-	0.4376	0.5803	2.1100	13.330	0.7468
BBB	0.8395	0.6552	0.2876	-	0.3709	0.3790	0.9105	0.6553
BB	1.6340	0.6732	0.4816	0.2186	-	0.3352	0.5151	0.6185
B	8.1140	2.7460	0.7948	0.5779	0.2728	-	0.3891	0.4470
CCC	7.4430	0	6.3890	0.8465	0.6324	0.3109	-	0.1762

Table 4.14: Model 4B.4 with shared sector frailties

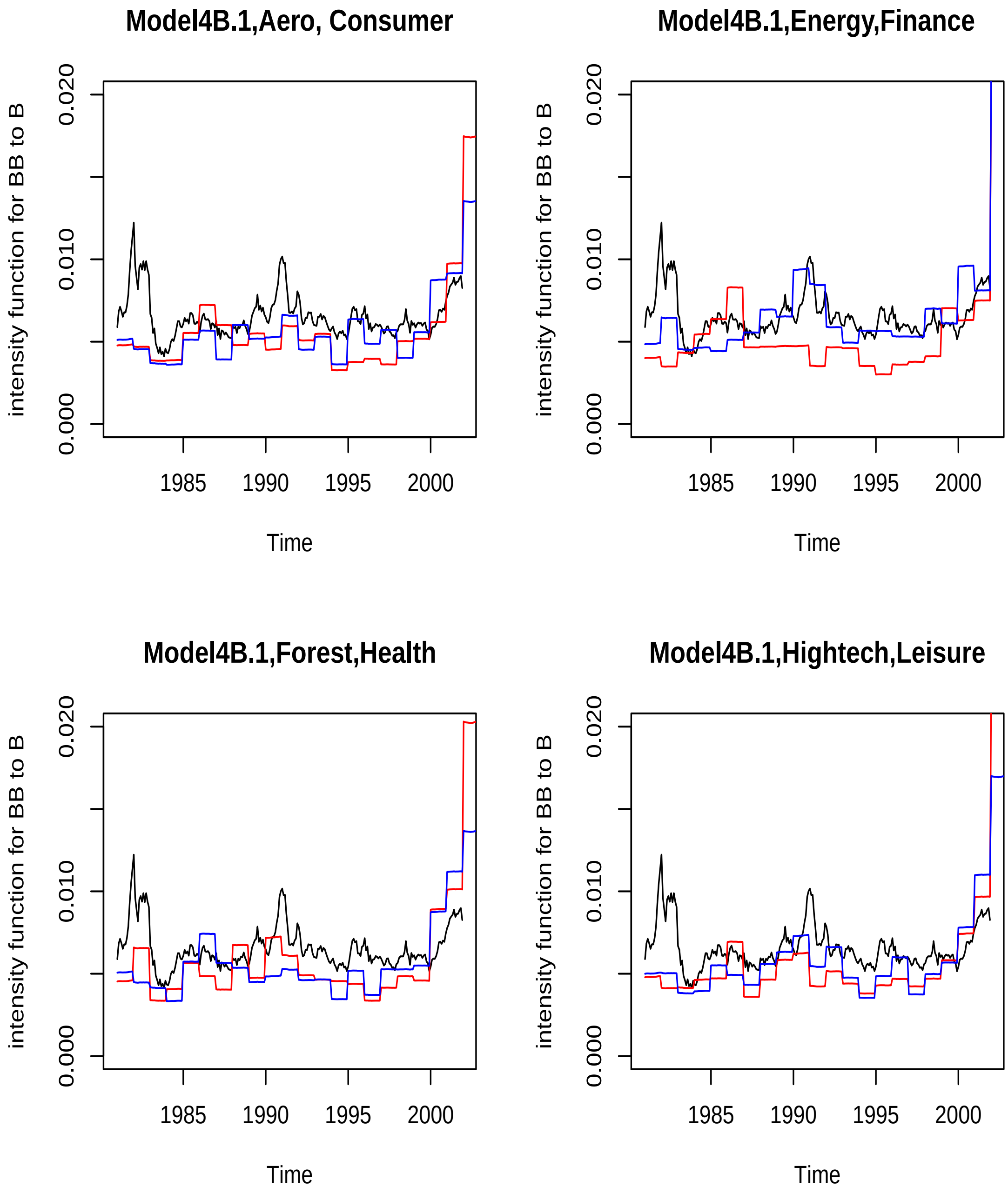

Figure 4.17: Monthly intensity for rating transition BB to B with Model 4B.1 (black line), Model 4B.4 (red and blue line)

Model 4B4: Posterior mean for γ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	-1.9410	-2.5530	-6.3410	-28.940	0	0	0
AA	-2.0880	-	-3.1810	-3.9600	-2.6570	-27.470	0.7132	0
A	1.3280	-1.6570	-	-3.2040	-3.7900	-11.210	-27.520	-2.4590
BBB	-2.5190	-3.4130	-2.0560	-	-2.6230	-2.2930	1.8570	-4.3710
BB	-4.9580	-0.8052	-2.7680	-1.4070	-	-2.3740	-3.3890	-4.3060
B	-26.060	-9.7190	-2.0030	-3.6820	-1.8920	-	-2.7430	-3.1170
CCC	-29.730	0	26.810	-1.5550	-3.7000	-2.0990	-	-1.2050

Model 4B4: Posterior standard deviation for γ								
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-	0.6933	0.9569	6.7040	18.720	0	0	0
AA	0.9235	-	0.5204	0.9424	2.4450	11.750	29.330	0
A	3.0250	0.4607	-	0.5101	0.9254	3.9160	40.270	1.6630
BBB	1.8700	1.6450	0.4129	-	0.4283	0.7649	3.1800	1.1880
BB	4.3150	1.9700	1.0950	0.3702	-	0.3929	0.8196	0.8636
B	20.390	6.900	2.1950	0.8965	0.3409	-	0.4450	0.5521
CCC	20.420	0	14.980	2.6730	1.2240	0.4465	-	0.2349

Table 4.15: Model 4B.4 with shared sector frailties

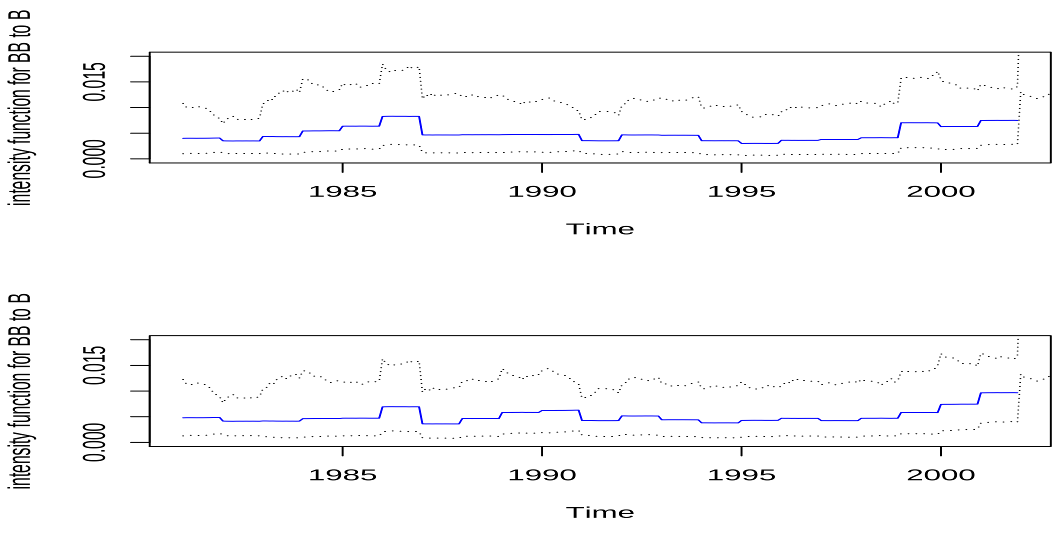

Figure 4.18: Monthly intensity for rating transition BB to B; posterior means with 95% credible intervals, the upper plot shows sector Energy and lower plot for sector Hightec

	No.Parameters	Dbar	Dhat	pD	DIC
Model 4B.1	88	4667310.00	4667230.00	82.354	4667390.00
Model 4B.2	158	4665410.00	4665270.00	133.185	4665540.00
Model 4B.3	160	4665400.00	4665270.00	130.515	4665530.00
Model 4B.4	320	4665070.00	4664850.00	226.707	4665300.00

Table 4.16: Deviance information criterion for four different frailty models

this case, monthly intensity are decided by yearly shared frailty (keep constant for 12 months) for each sector, therefore monthly intensities looks piece-wise constant[2]. Sector Aero and Consumer have strong co-movement while sector Energy and Finance are moving differently. Figure 4.18 shows the posterior mean of monthly intensity with 95% credible intervals for sector Energy and Hightec. For simplicity, we only pick up these two sectors to see the difference. We can clearly see difference between these two sectors, other sectors also have heterogeneity.

We use deviance information criterion (DIC) to compare these different models. For each model, the Gibbs sample was run for 50000 iterations, the first 25000 iterations were discarded, and the remaining 25000 iterations for each chain were used for analysis. The estimates for DIC and p_D for the four different models are presented in Table 4.16. The use of DIC is rather to compare different models than choose the true model. Model 4B.2 and Model 4B.3 have been introduced one level random effect and there are better than Model 4B.1. And Model 4B.3 has been introduced an AR(1) process, it is slightly better than Model 4B.2 from the DIC estimation. By introducing second level of random effect, Model 4B.4 improved compared with Model 4B.2 and Model 4B.3

[2]we are trying to show different kinds of results other than pick up the best results

4.4 Discussion

In this chapter, we extend our credit survival model from default risk to all transition risk. Two kinds of model were implemented, the first one is frailty model for credit rating transitions by numbers of levels (notches) which is a simpler case for credit migration data. It is easy to handle but sacrifice the accuracy for rating transition, therefore we finally model the actual rating transitions. Systematic portfolio risk is divided into observed fixed effects and unobserved random effects which were known as frailty in survival analysis to capture heterogeneity in migration analysis. We have shown heterogeneity of transition risk over time and industry sector. We can also shown heterogeneity for different countries if we extend our database to all the countries in Creditpro database. It is very ambitious and challenging to model the whole default and migration risk within one model since we have a very huge database, some parameters are estimated with quite large errors because of low numbers of transitions. With the limitation of time and resource got for research, the frailty model we used here can be extended by adding firm-specific covariates which is crucial for financial practitioners.

Chapter 5

Estimating default probabilities and transition matrices

Transition matrices are at the center of modern credit risk management. Transition matrices are widely used for risk management purpose, economic capital purpose and credit derivatives pricing. We have modelled credit risk using both GLMMs and survival analysis for default and transition risks. The default probabilities can be easily computed in GLMMs model. The transition matrices need to be calculated using the results in survival models. We have four aims in this chapter: First, we need to calculate transition matrices and default probabilities from intensities with survival model output. Second, we will show appropriate graphics of how transition rates change over time. Third, we will contrast with a stationary through-the-cycle estimate. Fourth, we will discuss any sectoral variation.

The modeling and estimation of transition matrices is an important issue because of the requirement of Basel III. Lando and Skodeberg (2002) give a review of different approaches to estimate migration matrices. Lando and Skodeberg (2002), and Christensen, Hansen and Lando (2004) address the issue of computing point and interval estimates for default probabilities with rare events, using a continuous time homogeneous Markov chain transition matrix. Some rare events like AAA to default will have no observation in the period of observation, but the transition rate is non-zero. There is also considerable evidence that the Markovian assumption for ratings transitions is unrealistic (Altman and Kao, 1992; Nickell, Perraudin and Varotto, 2000; Bangia et al., 2002; Frydman and Schuermann, 2005; Chava, Stefanescu and Turnbull, 2006). Jarrow, Lando and Turnbull (1997) make the distinction between implicit and explicit estimation of transition matrices. Implicit estimation refers to extracting transition matrices (including default probability) from market prices of risky zero-coupon bonds while explicit use historical transition information.

In this chapter, we will briefly introduce the time homogeneous hazard rate approach first. Then we develop a time-inhomogeneous model, assume the stochastic process is stationary within each month which is piece-wise constant. We will focus on the explicit methods and explore the time inhomogeneous hazard rate approach. We will use the output in previous chapters to calculate time-varying or time inhomogeneous transition and default rates.

5.1 Time homogeneous hazard rate approach

The cohort methods calculate the transition matrix by the ratio of total number of transitions from state h to j and the number of companies in state i during the observation period. An important consequence of this is that if the transition from h to j does not occur in the observation period, the estimate transition probability will be zero although it is non-zero. We will start with the time homogeneous hazard rate approach. The primary assumption of time homogeneous approach is that credit rating migrations are a homogenous Markov chain. The time homogeneous method (used by Lando and Skodeberg) is a method based on the assumption of a stationary continuous-time Markov chain. Let R_t denote a stochastic process taking values in $S = \{0, 1, ..., n\}$ at times $t = 0, 1, \cdots$. The set S defines rating state and (R_t) models the evolution of an obligor's rating over time. The stochastic process R_t is a Markov chain if for all $t \geq 1$ and all $h, r_0, r_1, \cdots, r_{t-2}, j \in S$

$$P(R_t = h | R_0 = r_0, R_1 = r_1, \cdots, R_{t-1} = j) = P(R_t = h | R_{t-1} = j)$$

This means conditional probabilities of rating transitions given an obligor's rating history depend only on the previous rating at the last time. The Markov chain is *stationary* if

$$P(R_t = h|R_{t-1} = j) = P(R_1 = h|R_0 = j)$$

for all $t \geq 1$ and all rating states h and j.

Following Lando & Skodeberg (2002), transition matrices can be described by a $k \times k$ generator or intensity matrix $\mathbf{\Lambda}$, and define $\boldsymbol{P}(t)$ is a $k \times k$ matrix of probabilities where hj^{th} element is the probability of migration from start state h to j at time period t: For a small time step δt we assume that the transition probability from rating h to j is given approximately by $\lambda_{hj}\delta t$. The probability of staying at rating h is given by $1 - \sum_{j \neq h} \lambda_{hj}\delta t$. Define a matrix Λ to have off-diagonal entries λ_{hj} and diagonal entries $-\sum_{j \neq h} \lambda_{hj}$, summarise these transition probabilities for a small time step in the matrix $I_{n+1} + \Lambda\delta t$. In the period $[0, t]$, let $\delta t = t/N$ for N steps. The matrix of transition probabilities $\boldsymbol{P}(t)$ can be approximated by

$$\boldsymbol{P}(t) \approx \left(I_{n+1} + \frac{\Lambda t}{N}\right)^N$$

which converges, as $N \to \infty$ to the matrix exponential of δt

$$\boldsymbol{P}(t) = \exp(\mathbf{\Lambda} t) \quad t \geq 0 \tag{5.1}$$

where the exponential is a matrix exponential, and $\mathbf{\Lambda}$ satisfy

$$\lambda_{hj} \geq 0 \quad h \neq j$$

$$\lambda_{hh} = -\sum_{h \neq j} \lambda_{hj} \tag{5.2}$$

λ_{hh} gives the diagonal elements which are chosen to ensure the rows sum to zero. λ_{hj} is obtained in previous chapters. An obligor remains in rating state h for an exponentially distributed amount of time with parameter $\sum_{h \neq j} \lambda_{hj}$. To estimate the element of the generator under time-homogenous assumption, we use maximum likelihood estimator:

$$\hat{\lambda}_{hj} = \frac{N_{hj}(T)}{\int_0^T Y_h(t)dt}$$

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	0.9938	0.0056	0.0005	0.0000	0.0001	0.0000	0.0000	0.0000
AA	0.0005	0.9925	0.0065	0.0004	0.0001	0.0001	0.0000	0.0000
A	0.0001	0.0017	0.9928	0.0001	0.0003	0.0050	0.0000	0.0000
BBB	0.0000	0.0002	0.0037	0.9907	0.0048	0.0005	0.0001	0.0001
BB	0.0000	0.0001	0.0004	0.0047	0.9853	0.0087	0.0006	0.0002
B	0.0000	0.0001	0.0002	0.0003	0.0043	0.9850	0.0075	0.0026
CCC	0.0001	0.0000	0.0003	0.0006	0.0012	0.0102	0.9347	0.0529
D	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

Table 5.1: The one month transition matrix calculated using the homogeneous hazard rate methods. The matrix was calculated for the period from Jan1980 to Dec2003

where $Y_i(s)$ is the number of obligors with rating h at time t.

We use the Standard & Poor's database CreditPro 6.6 over the period January 1981 to December 2003 for computing monthly transition probability. Using 5.1, the one month homogenous transition matrix are presented in Table 5.1:

5.2 Time non-homogeneous hazard rate approach

The homogeneous assumption means that the transition probability matrix for a given period only depend on the length of the period. However, the credit transition probability matrices also depend on the time period selected. Here we develop a time-inhomogeneous model, assume the stochastic process is stationary within each month which is piece-wise constant. This assumption is the same as in our modelling migration and default risk with survival models. For each month, we estimate a separate generator matrix and one month transition probability matrix based on the theory for homogeneous models in the previous section. We calculate one month transition probability matrix using 5.1. By composition, we get an estimated transition matrix between two arbitrary times t_1, t_2

$$\hat{\boldsymbol{P}}(t_1, t_2) = \prod_{t=t_1}^{t_2-1} \boldsymbol{P}(t) \tag{5.3}$$

where P_t stand for the matrix of transition probabilities in month t. Define $\boldsymbol{P}(t)$ is a $k \times k$ matrix of probabilities where hj^{th} element is the probability of migration from start state h to j at time period t. An one-year transition matrix is obtained by composition of 12 successive monthly transition matrices.

In the previous two chapters, we have used three different survival models for default and transition risk. We only consider default case for Model 3 and by numbers of levels (notches) for transition risk in Model 4A. Both of them are not suitable for computing transition matrices because only part of the transition risks are modelled in these two models. In Model 4B we have calibrate the frailty model for migration risk with actual transition types which cover all the transition types in our database. The S&P database has recorded 44 different transition types with 5 rating transitions have no observations from 1981-2003. We have discussed the calibration results in Chapter 4, the intensity function changes over time and shows the sectional variation. In this chapter, we will calculate the transition matrices with our output in Chapter 4 and show how the transition probability changes over time as well as the sectional variation for transition probability.

We have four different models fitting for transition risk for 23 years database, therefore we calculated 23 years transition matrices for four different models. Using 5.1, we get

$$\boldsymbol{P}_t = \exp(\boldsymbol{\Lambda}_t) \tag{5.4}$$

where $\boldsymbol{\Lambda}_t$ is the generator for month t, which can be defines as in the estimated model in Model 4B. For each element of $\boldsymbol{\Lambda}_t$ from rating state h to j, $\lambda_{hjt} = \exp(\boldsymbol{\eta}_{hj}^{'} \boldsymbol{z}_i(t) + \boldsymbol{\upsilon}_{hj} + \boldsymbol{\gamma}_{hj} b_{y(t)})$ where $y(t)$ gives the year period where t belongs to, time-dependent macroeconomic covariates $\boldsymbol{z}_t$, rating transition effect $\boldsymbol{\upsilon}$ and yearly frailty. We use the estimated results in Model 4B, then we get monthly transition probability matrix. As we fitted four models in the previous chapter, Model 4B.2 is chosen for showing

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	0.9827	0.0070	0.0038	0.0046	0.0016	0.0001	0.0001	0.0000
AA	0.0097	0.9223	0.0069	0.0073	0.0065	0.0151	0.0320	0.0003
A	0.0330	0.0211	0.8729	0.0164	0.0240	0.0229	0.0051	0.0046
BBB	0.0013	0.0041	0.0044	0.9737	0.0039	0.0051	0.0060	0.0015
BB	0.0054	0.0051	0.0063	0.0058	0.9630	0.0062	0.0062	0.0018
B	0.0020	0.0060	0.0084	0.0053	0.0070	0.9600	0.0063	0.0050
CCC	0.0001	0.0001	0.0031	0.0108	0.0046	0.0083	0.9610	0.0121
D	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

Table 5.2: The one month transition matrix calculated using the inhomogeneous methods for Dec2003

the calculation. The rest three models are the same. Table 5.2 shows the one month transition matrix for December 2003 which is different for every month.

In order to show how transition rates change over time, three different rating transition types are chosen and we will show their transition probabilities through-the-cycle. For generality, we choose rating transitions (BBB,A), (AA,A) and (BB,BB) which are stand for upgrade, downgrade and censor case separately. For simplicity, we only show these three transition types probabilities for Model 4B.2 because Model 4B.1 is a simple model among the models we fitted. The lines shows homogeneous hazard rate approach for three different rating transitions respectively. The credit transition probability for homogeneous model only depend on the length of time which means the monthly transition probability keeps constant over the observation period. While time inhomogeneous shows it changes over time. See figure 5.1: Transition probabilities changes through-the-cycle, especially we can find that upgrade and downgrade changes to different direction. Of course we find some noise around year 1990 where both downgrade and upgrade have relatively higher probabilities. Around year 1990 and 2000, we have a higher $\gamma_{AA\to A}b_t$. This result is very useful for bank's regulatory capital requirements and pricing credit derivatives.

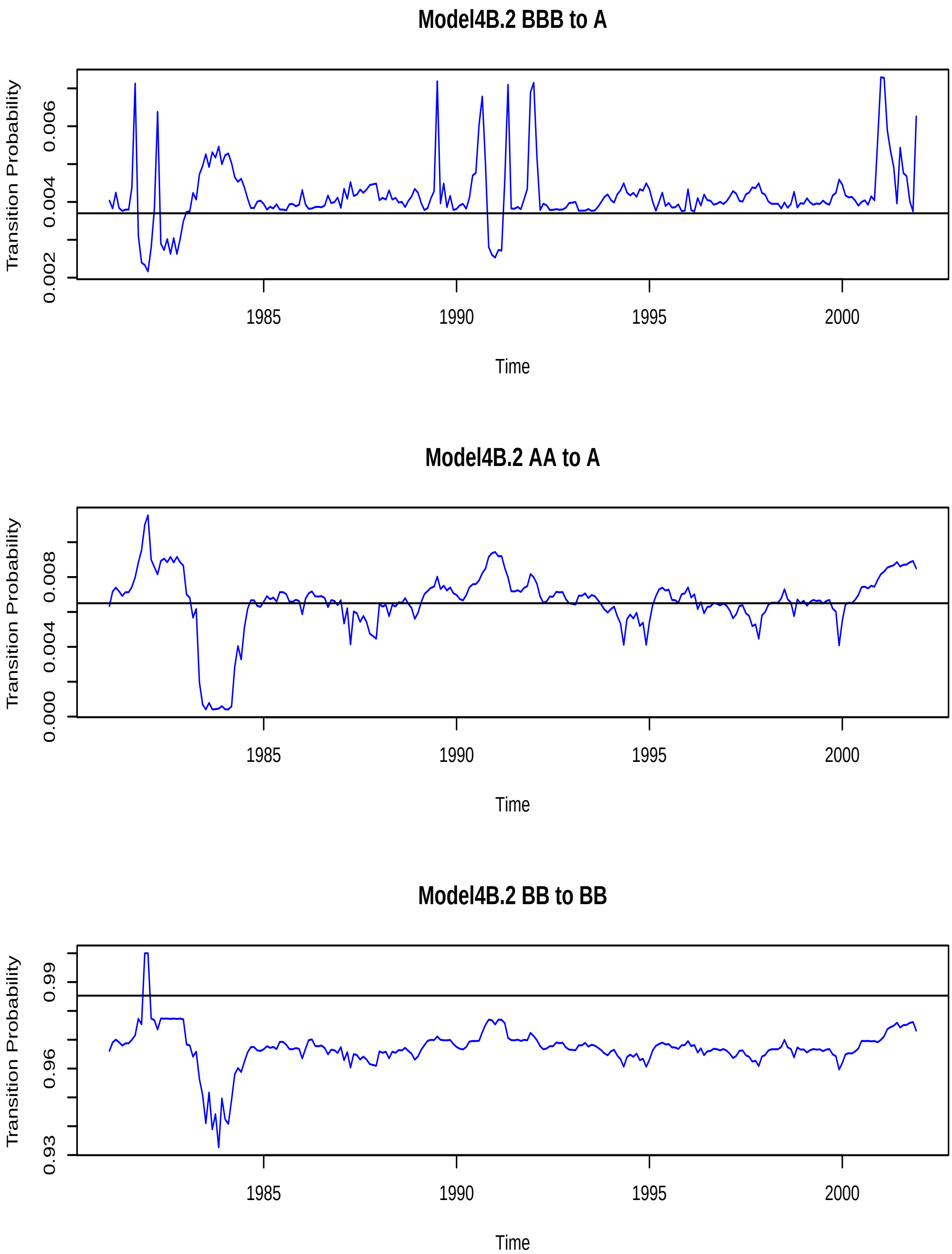

Figure 5.1: Monthly transition probabilities for three different rating transitions

Using 5.1, we get

$$\boldsymbol{P}_{ts} = \exp(\boldsymbol{\Lambda}_{ts}) \tag{5.5}$$

where $\boldsymbol{\Lambda}_{ts}$ is the generator for month t, which can be defined as the estimated model in Model 4B.4. For each element of $\boldsymbol{\Lambda}_t$ from rating state h to j, $\lambda_{hjts} = \exp(\boldsymbol{\eta}^{'}_{hj}\boldsymbol{z}(t) + \boldsymbol{\upsilon}_{hj} + \boldsymbol{\gamma}_{hj} b_{y(t),s})$. It depends on time-dependent macroeconomic covariates $\boldsymbol{z}_t$, rating transition effect and two levels of shared frailties of yearly and sector.

With the output of Model 4B.4, we plot the transition probability for different sectors and capture the heterogeneity for different sectors which cannot be explained by the one-level yearly shared frailty model. We choose transition (BB,B) for Model 4B.4 as an example. The black line shows one-level shared frailty model and red line for section Aero Energy Forest and Hightech and blue line for rest of the section. Finance sector has higher default probability around 1990 while other sectors relatively low. From figure 5.2, we can clearly see the sectoral variation.

5.3 Summary

The output of Model 4B in Chapter 4 has been used for calculating matrices of transition probability in this chapter. While transition rates in the time homogeneous model only depend on the length of time in a rating state, in the time inhomogeneous model the transition rates clearly change over time. They show considerable variation around the stationary transition rates. We also have shown there is sectional variation between transition rates. Our model can be used to forecast transition rates in the future. We note that this also requires the forecasting of the macroeconomic covariate CFNAIMA3.

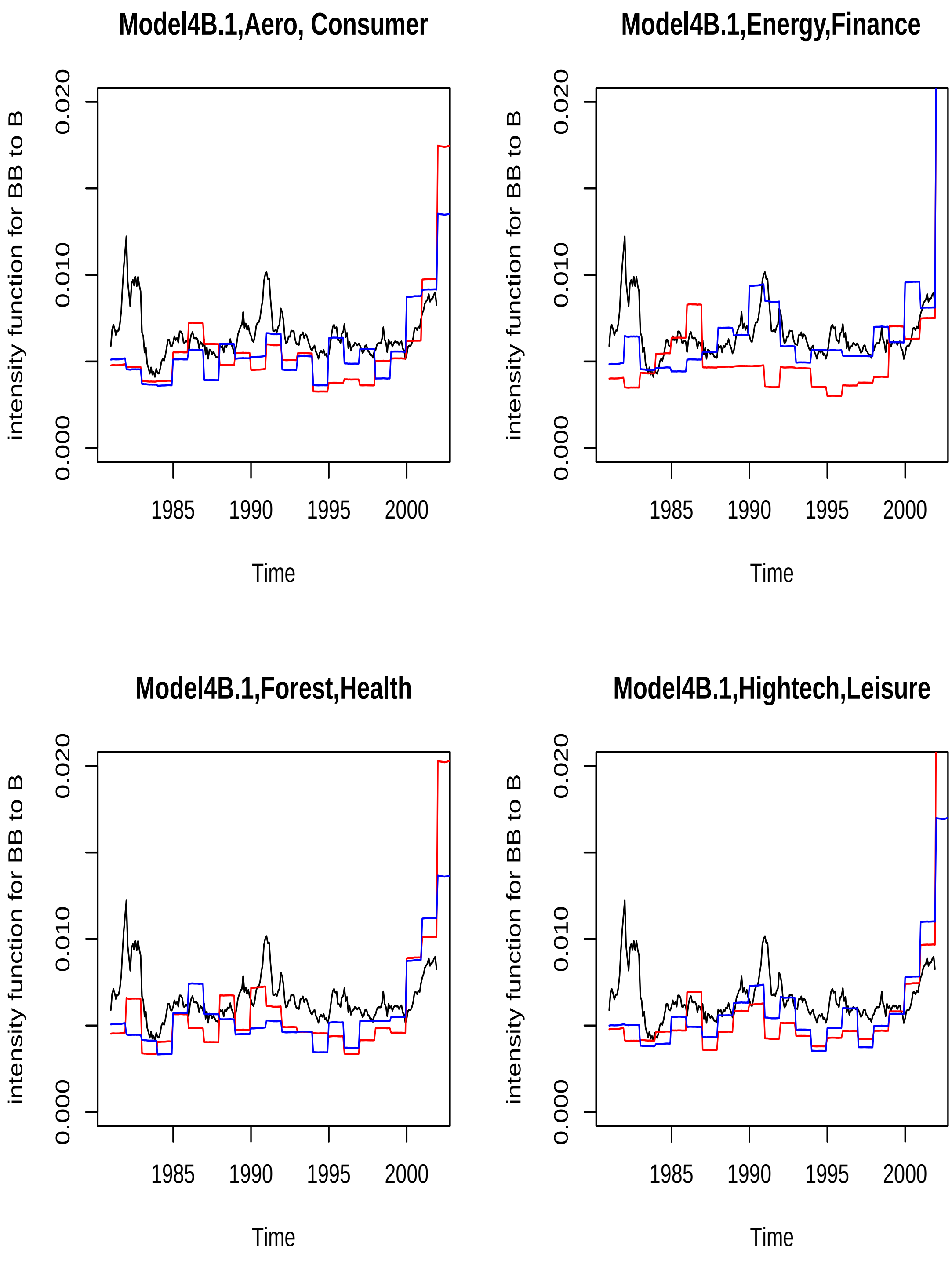

Figure 5.2: Monthly transition probabilities for 8 different industries

Chapter 6

Summary and conclusions

To manage credit risk effectively we need to identify sources of heterogeneity in default risk and rating migration risk. The most fundamental determinant of heterogeneity in credit portfolio is the credit quality or rating of the obligors; the probability of default differs across different rating classes. Basel III introduced more information than credit quality to explain heterogeneity.

In this thesis, we have explored some important additional sources of heterogeneity. First, the default risk varies due to time-varying macroeconomic covariates. Among a list of macroeconomic covariates we find a smoothed version of the Chicago Fed National Activity Index (CFNAIMA3) is the best observed macroeconomic covariates to describe credit quality changes. Second, the default and transition risk also relate to industry sectors and countries. Different industry sectors show different default and transition probability and obligors sharing industry sector show higher levels of default correlation than others. This issue proposed by BCBS (2002) as concentration risk which violate single-factor models assumption.

We use two different kinds of statistical models - GLMMs and survival models - for modelling default and transition risk. In both models, our empirical study using CreditPro data shows that the most predictive credit model includes observed macroeconomic covairates CFNAIMA3, random unobserved factors like time and industry sectors. This implies that the rating transition matrix depends on the state of the economy, time and industry sectors. We consider generalised linear mixed model (GLMMs) for default count data where fixed effect account for observed factor risk and random effect for unobserved factor risk. GLMMs capture heterogeneity in time and across industry sectors in our research. With the consideration of industry-specific factors, we find the heterogeneity of industry sectors are one of the source of default risk. However, we are not only interested in the number of companies credit rating transfer but also the length of time for obligors stay in. Therefore we use time-to-event analysis to model default risk. Andersen & Gill model generalize Cox proportional hazard model to allow time-varying covariates using counting process formulation. The frailty in survival models captures time and sector heterogeneity with unobserved random variables. We introduce the statistical model which was used in medical research by Manda & Meyer into credit risk modelling for the first time. We show an empirical study on credit rating default risk with survival analysis before moving to transition risks model. The default risk model with survival analysis shows business cycle through the time period. The frailty captures unobserved random effects. We use two levels of frailty, corresponding to time and sector, to estimate intensity function of default risk and in some models we include serial dependence for time random effect.

We extend the model to all transition risk and model numbers of levels (notches) and actual rating transitions in a survival analysis framework. With the results of transition risk, we develop a time inhomogeneous model with piecewise constant models to calculate transition matrix probabilities. We compute the transition matrices from the intensity output in transition risk model and show how transition rate change over time and display sectoral variation. With the output from the transition risk model, we can forecast transition probabilities with the forecast of macroeconomic covariates. We try to model transition risk within one model because of our ambition. The estimation results of transition with less events are not as good as the transition with plenty of events in our observation time period. The public database in credit risk has been improved in the last few years. We show an applicable way of modelling transition risk and forecasting transition matrices with our models.

Bibliography

[1] Altman, E. (1968). Financial ratios: discriminant analysis, and the prediction of corporate bankruptcy. *Journal of Finance* **23,** 589-609.

[2] Andersen. PK, Gill. RD (1982). Cox regression models for counting processes. *The Annals of Statistics* **10,** 1100-1120.

[3] Andersen, P.K., Borgan, O., Gill, R.D. and Keiding, N. (1993) *Statistical Models Based on Counting Processes.* Springer-Verlag.

[4] Bangia, A., F. X. Diebold, A. Kronimus, C. Schagen, and T. Schuermann (2002). Ratings migration and the business cycle, with application to credit portfolio stress testing. *Jornal of Banking and Finance* **26,** 445-474.

[5] Basel Committee on Banking Supervision (BCBS) (2002, October). Quantitative impact, Study 3, Technical guidance. Bank of International Settlements.

[6] Basel Committee on Banking Supervision (BCBS) (2004). Basel II: International convergence of capital measurement and capital standards: a revised framework, Bank of International Settlements.

[7] Basel Committee on Banking Supervision (BCBS) (2005, May). Studies on validation of internal rating systems. *Working paper No. 14* Bank of International Settlements.

[8] Bellotti,T., and J. Crook (2009). Credit scoring with macroeconomic variables using survival analysis. *Journal of the Operational Research Society* **60**, 1699-1707.

[9] Breslow,N. and D. Clayton (1993). Approximate inference in generalized linea mixed models *Jonural of American Statistical Association* **88**, 9-25.

[10] Carlin,B.P. and T.A. Louis (2000). *Bayes and Empirical Bayes Methods for Data Analysis* (2nd ed.). Chapman & Hall/CRC, Boca Raton

[11] Catalina,S.,R. Tunaru and S. Turnbull (2009). The credit rating process and estimation of transition probabilities: a bayesian Approach, *Journal of Empirical Finance*, **16**, 216-234.

[12] Chava, S., and R. Jarrow, (2004). Bankruptcy prediction with industry effects, *Review of Finance* **8(4)**, 537-569.

[13] Chava,S., C. Stefanescu., and S. Turnbull., (2008). Modelling expected loss *Working paper*, Bauer College of Business, Univeristy of Houston.

[14] Clayton. D (1991). A Monte Carlo method for Bayesian inference in frailty models *Biometrics* **47,** 467-485.

[15] Clayton, D (1996). Generalized linear mixed models. In: Gilks,W., Richardson,S., Spiegelhalter,D.(Eds.), *Markov Chain Monte Carlo in Practice*. Chapman& Hall, London, 275-301.

[16] Couderc. F, Renault, O. (2005). Time-to-default: Life cycle, global and industry cycle impacts, *Working paper*, University of Geneva.

[17] Crouhy, M., Galai, D., and Mark, R. (2000). A comparative analysis of current credit risk models. *Journal of Banking Finance* **24**, 59-117.

[18] Delloye, M., Fermanian J.D., and Sbai, M. (2006), Estimation of a reduced-form credit portfolio model and extensions to dynamic frailties. *Working Paper*, CREST

[19] Egloff, D., Leippold M., and Vanini, P. (2007), A simple model for credit contagion. *Journal of Banking and Finance*, **31**, 2475-2492.

[20] Frey,R. and A.J.McNeil (2003). Dependent defaults in models of portfolio risk. *Journal of Risk* **6(1)**, 59-92.

[21] Gagliardini, P. and C. Gourieroux (2005a). Migration correlation: Definition and efficient estimation. *Journal of Banking Finance* **29**, 865-894.

[22] Gagliardini, P. and C. Gourieroux (2005b). Stochastic migration models with application to corporate risk. *Journal of Financial Econometrics* **3(2)**, 188-226.

[23] Gordy, M., (2000). A comparative anatomy of credit risk models, *Journal of Banking and Finance* **24(1)**, 119-149.

[24] Gordy, M., (2003). A risk-factor model foundation for ratings-based bank capital rules, *Journal of Financial Intermediation*, **12**, 199-232.

[25] Gordy, M. B. and E. Heitfield (2002). Estimating default correlations from short panels of credit rating performance data. Technical Report, Federal Reserve Board.

[26] Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Springer-Verlag, New York.

[27] Hsiao, C. (2003). *Analysis of Panel Data* (2nd ed.). Cambridge University Press, Cambridge.

[28] Hu,Y.-T., R. Kiesel, and W. Perraudin (2002). The estimation of transition matrices for sovereign credit ratings. *Journal of Banking and Finance* **26,** 1383-1406.

[29] Jafry, Y. and T. Schuermann (2004). Measurement, estimation and comparison of credit migration matrices. *Journal of Banking and Finance* **28**, 2603-2639.

[30] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

[31] Kadam A, Lenk, P. (2008). Bayesian inference for issuer heterogeneity in credit ratings migration, *Journal of Banking and Finance*, **32(10)**, 2267-2274.

[32] Kavvathas, D. (2000). Estimating credit rating transition probabilities for corporate bond, *Working paper*, University of Chicago.

[33] Koopman,S.J., A. Lucas, and O. Klaassen (2005). Empirical credit cycles and capital buffer formation. *Journal of Banking and Finance* **29,** 3159-3179.

[34] Koopman, S. J., A. Lucas, and A. Monteiro (2005). The multi-state latent factor intensity model for credit rating transitions. *Journal of Econometrics* **142**, 399-424.

[35] Koopman, S. J., A. Lucas, and B. Schwaab (2012). Dynamic factor models with macro, frailty, and industry effects for U.S. default counts: the credit crisis of 2008. *Journal of Business and Economic Statistics* **30**, 521-532.

[36] Lando, D. (2004). *Credit RiskModeling: Theory and Applications.* Princeton University Press, Oxford.

[37] Lando, D. and T. Skodeberg (2002). Analyzing rating transitions and rating drift with continuous observations. *Journal of Banking Finance* **26**, 423-444.

[38] Manda, S.O.M., Meyer, R. (2005). Bayesian inference for recurrent events data using time-dependent frailty, *Statistics in Medicine* **24**, 1263-1274.

[39] McCullagh, P. and J. A. Nelder (1989). *Generalized Linear Models* (2nd ed.). Chapman & Hall, London.

[40] McNeil, A.J., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management.* Princeton University Press, Princeton.

[41] McNeil, A.J., and J. Wendin, (2006) Dependent credit migrations. *Journal of Credit Risk*, **2(3)**, 87-114.

[42] McNeil, A.J., and J. Wendin, (2007). Bayesian inference for generalized linear mixed models of portfolio credit risk. *Journal of Empirical Finance* **14,** 131-149.

[43] Metayer,B. (2006). Shared Frailty Model for Rating Transitions, *Working paper*, University of Zürich.

[44] Nickell, P., W. Perraudin, and S. Varotto (2000). Stability of rating transitions. *Journal of Banking Finance* **24,** 203-227.

[45] Robert, C. P. and G. Casella (1999). *Monte Carlo Statistical Methods.* Springer-Verlag, New York.

[46] Rösch, D. (2005) An Empirical Comparison of Default Risk Forecasts from Alternative Credit Rating Philosophies. *International Journal of Forecasting*, **21**, 37-51.

[47] Wendin,J., (2006). Bayesian Methods in Portfolio Credit Risk Management. Ph.D. thesis, *Department of Mathematics, ETH Zürich.*

[48] Wilson, T. (1997). Portfolio credit risk I and II. *Risk* 10(9, 10).

[49] Sbai,M. M. Delloye, J-D. Fermanian.(2006). Dynamic frailties and credit portfolio modelling. *Risk* **19(10),** 100-105.

[50] Servigny, A.D., O. Renault (2004). *Measuring and Managing Credit Risk.* McGraw Hill Professional.

[51] Shumway, T. (2001). Forcsting Bankruptcy More Accurately: A simple hazard model. *Journal of Business* **74**, 101-124.

[52] Spiegelhalter DJ, Thomas A, Best NG, Gills WR. (2003). BUGS: Bayesian Inference Using Gibbs Sampling, Version 1.4. Medical Research Council Biostatistics Unit, Cambridge Univeristy.

[53] Zmijewski, M.E. (1984), Methodological issues related to the estimation of financial distress prediction models, *Journal of Accounting Research*, **22**, 59-82.